(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 January 2004 (08.01.2004)

PCT

(10) International Publication Number
WO 2004/003013 A1

(51) International Patent Classification[7]: C07K 14/415

(21) International Application Number:
PCT/EP2003/006757

(22) International Filing Date: 26 June 2003 (26.06.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0214896.3 27 June 2002 (27.06.2002) GB

(71) Applicant *(for all designated States except US)*: **NOVARTIS FORSCHUNGSSTIFTUNG, ZWEIGNIEDERLASSUNG FRIEDRICH MIESCHER INSTITUTE FOR BIOMEDICAL RESEARCH** [CH/CH]; Maulbeerstrasse 66, CH-4058 Basel (CH).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **FRITSCH, Olivier** [FR/FR]; 9, Rue des Artisans, F-68000 Colmar (FR). **HOHN, Barbara** [AT/CH]; Hangstrasse 35, CH-4144 Arlesheim (CH). **LUCHT, Jan, Martin** [DE/CH]; Unterer Batterieweg 113, CH-4059 Basel (CH).

(74) Common Representative: **NOVARTIS FORSCHUNGSSTIFTUNG, ZWEIGNIEDERLASSUNG FRIEDRICH MIESCHER INSTITUTE FOR BIOMEDICAL RESEARCH**; Maulbeerstrasse 66, CH-4058 Basel (CH).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2004/003013 A1

(54) Title: GENE FOR INCREASED SOMATIC RECOMBINATION

(57) Abstract: The present invention relates to nucleic acids encoding polypeptides involved in homologous recombination, as well as vectors and host cells comprising the nucleic acids and polypeptides encoded by the nucleic acids. Also provided are methods for inducing somatic and/or meiotic homologous recombination in a cell, comprising modulating the expression or properties of one or more gene products selected from the group consisting of AtIno80, At3g57300, Rvb1 (At5g22330), Rvb21 (At5g67630), Rvb22 (At3g49830), At3g57290, AtArp5.1 (At3g12380), AtArp5.2 (At5g56180), AtArp5.3 (At3g60830) and AtArp8 (At5g43500), their homologues, fragments or derivatives. In particular, the methods can be used to increase gene targeting.

Gene for increased somatic recombination

TECHNICAL FIELD

The present invention relates to DNA that encodes proteins that control somatic recombination, in particular in plants.

BACKGROUND

Cells of all organisms have evolved a series of DNA repair pathways that counteract the deleterious effects of DNA damage and are triggered by intricate signal cascades. Homologous recombination in plants stabilizes the genome by repairing damaged chromosomes simultaneously generating genetic variability through the creation of new genes and new genetic linkages. Repair of DNA damage by recombination is particularly significant for cells under exogenous and endogenous genotoxic stress because of its potential to remove a wide range of DNA lesions. The current understanding of genetic and molecular components underlying meiotic and somatic recombination and DNA repair in plants is limited. To be able to modify or improve DNA repair using gene technology it is necessary to identify key proteins involved in said pathways or cascades.

The precise manipulation of the genome of higher plants is still a major challenge for plant genetic engineering. Some advances have been made recently for the creation of point mutations at predetermined positions by chimeric RNA/DNA oligonucleotides (Beetham et al. 1999, Hohn & Puchta 1999, Zhu et al. 1999, Kipp et al. 2000, Zhu et al. 2000). However, the targeted insertion of longer stretches of DNA sequence at any desired location ("knock-in") or the replacement of predetermined plant genomic sequences by heterologous DNA ("knock-out) via homologous recombination is at present not possible as a routine technique (Mengiste & Paszkowski 1999, Puchta 2002).

Few reports have appeared in the literature that describe successful "gene targeting" in higher plants (Paszkowski et al. 1988, Lee et al. 1990, Offringa et al. 1990, Miao & Lam 1995, Kempin et al. 1997, Hanin et al. 2001), but the reported absolute numbers and relative

frequencies of the desired events were very low. Indeed, the main problem for "gene targeting" experiments is the low frequency of the desired homologous recombination events - 10^{-3} to 10^{-5} (Hohn & Puchta 1999, Mengiste & Paszkowski 1999) - relative to illegitimate recombination/integration events.

Various attempts of increasing the low relative frequency of targeted homologous recombination events, by improved selection schemes ("positive-negative selection") or by providing extended regions of sequence homology, were not successful (Thykjaer et al. 1997, Gallego et al. 1999). One promising strategy to facilitate gene targeting in higher plants would be to shift the balance between illegitimate and homologous recombination events towards the latter, by facilitating homologous recombination events in plants by genetic manipulation (Gherbi et al. 2001).

One approach described in the literature is the expression in plants of heterologous proteins known to be involved in homologous recombination. Overproduction of the bacterial resolvase RuvC was shown to increase somatic inter-and intra-chromosomal recombination, as well as extrachromosomal recombination (Shalev et al. 1999), but no gene targeting studies were reported yet with this system. Expression of the bacterial RecA protein had similar effects (Reiss et al. 1996, Reiss et al. 1997), but subsequent experiments did not show an increase of gene targeting events (Reiss et al. 2000). So far, it is not clear whether heterologous proteins can successfully interact with the plant recombination machinery to affect the outcome of the recombination events required for gene targeting. In addition, these foreign proteins might have undesired side effects in plants.

An alternative approach is to rely on endogenous plant genes to influence the frequency of homologous recombination events. So far, indirect approaches have been reported to isolate plant genes involved in recombination. The cloning of plant orthologs to recombination and repair genes from other species was reported (Klimyuk & Jones 1997, Doutriaux et al. 1998, Hartung & Puchta 1999, Gallego et al. 2000, Lin et al. 2000), but so far the importance of these genes for recombination in plants has only been evaluated for the RAD50 homologue (Gherbi et a., 2001). Functional screens have been carried out to identify plant mutants hypersensitive to genotoxic treatments (Davies et al. 1994, Jenkins et al. 1995, Jiang et al. 1997, Masson et al. 1997, Albinsky et al. 1999, Mengiste et al. 1999). Since recombination is an important mechanism for DNA repair, some of these mutants might be affected in their

recombination behavior. This was experimentally demonstrated for some X-ray hypersensitive Arabidopsis mutants that also showed reduced levels of somatic recombination (Masson & Paszkowski 1997), although the affected gene has not been isolated. Recently, a DNA damage hypersensitive Arabidopsis mutant was isolated from a T-DNA tagged population, the affected gene (MIM) was cloned and shown to encode an SMC (Structural Maintenance of Chromatin) protein. Since the *mim* mutant showed decreased frequencies of somatic recombination, MIM seems be involved in some aspect of somatic recombination (Mengiste et al. 1999). Also in tobacco a hyperrecombinogenic mutant was isolated (Gorbunova et al. 2000). However, the gene affected could not be isolated so far.

Previously, a genetic system was described to study somatic homologous recombination between repeated sequences in whole plants (Swoboda et al. 1994, Puchta et al. 1995a, Puchta et al. 1995b). Briefly, a transgene carrying two non-functional halves of the β-glucuronidase reporter gene sharing a stretch of sequence identity serves as a reporter construct. Homologous recombination between the repeated sequences results in the restoration of a functional reporter gene. Such events were detected by a sensitive histochemical assay, and confirmed by Southern blotting. This assay is destructive, since the staining procedure is lethal, so that direct isolation of mutants is difficult.

There is a need in the art to identify genes that increase somatic recombination and this invention meets that need.

RELATED LITERATURE

1) Aguilera, A. & Klein, H.L. Genetic control of intrachromosomal recombination in *Saccharomyces cerevisisae.* I. Isolation and genetic characterization of hyper-recombination mutations. Genetics 119: 779-790, 1985.

2) Albinsky, D. et al. Plant responses to genotoxic stress are linked to an ABA/salinity signaling pathway. Plant J 17:73-82, 1999.

3) Aravind, L. et al. Conserved domains in DNA repair proteins and evolution of repair systems. Nucl Acids Res 27:1223-1242, 1999.

4) Beetham, P.R. et al. A tool for functional plant genomics: chimeric RNA/DNA oligonucleotides cause in vivo gene-specific mutations. Proc Natl Acad Sci USA 96: 8774-8778, 1999.

5) Cho, S.-G. et al. TIP49b, a regulator of activating transcription factor 2 response to stress and DNA damage. Mol Cell Biol 21:8398-8413, 2001.

6) Clough, S.J. & Bent, A.F. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J 16:735-743, 1998.

7) Davies, C. et al. Isolation of Arabidopsis thaliana mutants hypersensitive to gamma radiation. Mol Gen Genet 243:660-665, 1994.

8) Davis, J.L. et al. A presumptive helicase (MOT1 gene product) affects gene expression and is required for viability in the yeast Saccharomyces cerevisiae. Mol Cell Biol 12:1879-1892, 1992.

9) Dilkes, B.P. & Feldmann, K.A. Cloning genes from T-DNA tagged mutants. Methods Mol Biol 82:339-351, 1998.

10) Doutriaux, M.-P. et al. Isolation and characterisation of the RAD51 and DMC1 homologs from Arabidopsis thaliana. Mol Gen Genet 257:283-291, 1998.

11) Ebbert, R. et al. The product of the SNF2/SWI2 paralogue INO80 of Saccharomyces cerevisiae required for efficient expression of various yeast structural genes is part of a high-molecular-weight protein complex. Mol Microbiol 32:741-751, 1999.

12) Emery, H.S. et al. Sequence of RAD54, a Saccharomyces cerevisiae gene involved in recombination and repair. Gene 104:103-106, 1991.

13) Essers, J. et al. Homologous and non-homologous recombination differentially affect DNA damage repair in mice. EMBO J 19:1703-1710, 2000.

14) Fang, R.X. et al. Multiple cis regulatory elements for maximal expression of the cauliflower mosaic virus 35S promoter in transgenic plants. Plant Cell 1:141-150, 1989.

15) Fridborg, I. et al. The Arabidopsis dwarf mutant shi exhibits reduced gibberellin responses conferred by overexpression of a new putative zinc finger protein. Plant Cell 11:1019-1032, 1999.

16) Gallego, F. et al. AtRAD1, a plant homologue of human and yeast nucleotide excision repair endonucleases, is involved in dark repair of UV damages and recombination. Plant J 21:507-518, 2000.

17) Gallego, M.E. et al. Positive-negative selection and T-DNA stability in Arabidopsis transformation. Plant Mol Biol 39:83-93, 1999.

18) Galli, A. et al. Characterization of the hyperrecombination phenotype of the *pol3-t* mutation of *Saccharomyces cerevisiae*. Genetics 164: 65-79, 2003.

19) Gamborg, O.L. et al. Nutrient requirement of suspension cultures of soybean root cells. Exp Cell Res 50:151-158, 1968.

20) Gherbi, H. et al. Homologous recombination in planta is stimulated in the absence of Rad50. EMBO Rep 2:287-291, 2001.

21) Gorbalenya, A.E. & Koonin, E.V. Helicases: amino acid sequence comparisons and structure-function relationships. Curr Opin Struct Biol 3:419-429, 1993.

22) Gorbunova, V. et al. A new hyperrecombinogenic mutant of Nicotiana tabacum. Plant J 24:601-611, 2000.

23) Guerineau, F. et al. Sulfonamide resistance gene for plant transformation. Plant Mol Biol 15:127-136, 1990.

24) Hanin, M. et al. Gene targeting in Arabidopsis. Plant J 28:671-677, 2001.

25) Hartung, F. & Puchta, H. Isolation of the complete cDNA of the Mre11 homologue of Arabidopsis (Accession No. AJ243822) indicates conservation of DNA recombination mechanisms between plants and other eucaryotes. (PGR99-132). Plant Physiol 121: 312, 1999.

26) Hayashi, H. et al. Activation of a plant gene by T-DNA tagging: auxinindependent growth in vitro [retracted by Schell J. Science 284:1275, 1999]. Science 258:1350-1353, 1992.

27) Hohn, B. & Puchta, H. Gene therapy in plants. Proc Natl Acad Sci USA 96:8321-8323, 1999.

28) Ikura, T. et al. Involvement of the TIP60 histone acetylase complex in DNA repair and apoptosis. Cell 102:463-473, 2000.

29) Jenkins, M.E. et al. Radiation-sensitive mutants of Arabidopsis thaliana. Genetics 140: 725-732, 1995.

30) Jelesko J.G. et al. Rare germinal unequal crossing-over leading to recombinant gene formation and gene duplication in Arabidopsis thaliana. Proc Natl Acad Sci USA 96:10302-10307, 1999.

31) Jiang, C.-Z. et al. UV- and gamma-radiation sensitive mutants of Arabidopsis thaliana. Genetics 147:1401-1409, 1997.

32) Jonsson, Z.O. et al. Rvb1p and Rvb2p are essential components of a chromatin remodeling complex that regulates transcription of over 5% of yeast genes. J Biol Chem 276:16279-16288, 2001.

33) Kakimoto, T. CKI1, a histidine kinase homolog implicated in cytokinin signal transduction. Science 274:9825, 1996.

34) Kanemaki, M. et al. TIP49b, a new RuvB-like DNA helicase, is induced in a complex together with another RuvB-like DNA helicase, TIP49a. J Biol Chem 274:22437-22444,

1999.

35) Kardailsky, I. et al. Activation tagging of the floral inducer FT. Science 286:1962-1965, 1999.

36) Kempin, S.A. et al. Targeted disruption in Arabidopsis. Nature 389:802-803, 1997.

37) Kipp, P.B. et al. Gene targeting in plants via site-directed mutagenesis. Methods Mol Biol 133:213-221, 2000.

38) Klimyuk, V.I. & Jones, J.D. AtDMC1, the Arabidopsis homologue of the yeast DMC1 gene: characterization, transposon-induced allelic variation and meiosis- associated expression. Plant J 11:1-14, 1997.

39) Laurent, B.C. et al. Functional interdependence of the yeast SNF2, SNF5, and SNF6 proteins in transcriptional activation. Proc Natl Acad Sci USA 88:2687-2691, 1991.

40) Lee, K.Y. et al. Homologous recombination in plant cells after Agrobacterium-mediated transformation. Plant Cell 2:415-425, 1990.

41) Liu, Z. et al. Repair of UV damage in plants by nucleotide excision repair: Arabidopsis UVH1 DNA repair gene is a homolog of Saccharomyces cerevisiae Rad1. Plant J 21:519-528, 2000.

42) Masson, J.E. et al. Mutants of Arabidopsis thaliana hypersensitive to DNA-damaging treatments. Genetics 146:401-407, 1997.

43) Masson, J.E. & Paszkowski, J. Arabidopsis thaliana mutants altered in homologous recombination. Proc Natl Acad Sci USA 94:11731-11735, 1997.

44) Mathur, J. et al. Gene identification with sequenced T-DNA tags generated by transformation of Arabidopsis cell suspension. Plant J 13:707-716, 1998.

45) Mayerhofer, R. et al. T-DNA integration: a mode of illegitimate recombination in plants. EMBO J 10:697-704, 1991.

46) Mengiste, T. & Paszkowski, J. Prospects for the precise engineering of plant genomes by homologous recombination. Biol Chem 380:749-758, 1999.

47) Mengiste, T. et al. An SMC-like protein is required for efficient homologous recombination in Arabidopsis. EMBO J 18:4505-4512, 1999.

48) Miao, Z.H. & Lam, E. Targeted disruption of the TGA3 locus in Arabidopsis thaliana. Plant J 7:359-365, 1995.

49) Michelet, B. & Chua, N.-H. Improvement of Arabidopsis mutant screens based on luciferase imaging in planta. Plant Mol Biol Rep 14:320-329, 1996.

50) Millar, A.J. Circadian clock mutants in Arabidopsis identified by luciferase imaging. Science 267:1161-1163, 1995a.

51) Millar, A.J. et al. Firefly luciferase as a reporter of regulated gene expression in higher plants. Plant Mol Biol Rep 10:324-327, 1992.

52) Millar, A.J. et al. The regulation of circadian period by phototransduction pathways in Arabidopsis. Science 267:1163-1166, 1995b.

53) Muchardt, C. & Yaniv, M. ATP-dependent chromatin remodelling: SWI/SNF and Co. are on the job. J Mol Biol 293:187-98, 1999.

54) Muris, D.F. et al. Isolation of the Schizosaccharomyces pombe RAD54 homologue, rhp54+, a gene involved in the repair of radiation damage and replication fidelity. J Cell Sci 109:73-81, 1996.

55) Nacry, P. et al. Major chromosomal rearrangements induced by T-DNA transformation in Arabidopsis. Genetics 149:641-650, 1998.

56) Offringa, R. et al. Extrachromosomal homologous recombination and gene targeting in plant cells after Agrobacterium mediated transformation. EMBO J 9:3077-3084, 1990.

57) Paszkowski, J. et al. Gene targeting in plants. EMBO J 7:4021-4026, 1988.

58) Pruitt, R.E. & Meyerowitz, E.M. Characterization of the genome of Arabidopsis thaliana. J Mol Biol 187:169-183, 1986.

59) Puchta, H. et al. Somatic intrachromosomal homologous recombination events in populations of plant siblings. Plant Mol Biol 28:281-292, 1995a.

60) Puchta, H. et al. Induction of intrachromosomal homologous recombination in whole plants. Plant J 7:203-210, 1995b.

61) Puchta, H. Gene replacement by homologous recombination in plants. Plant Mol Biol 48:173-182, 2002.

62) Reiss, B. et al. RecA protein stimulates homologous recombination in plants. Proc Natl Acad Sci USA 93:3094-3098, 1996.

63) Reiss, B. et al. Targeting of a functional Escherichia coli RecA protein to the nucleus of plant cells. Mol Gen Genet 253:695-702, 1997.

64) Reiss, B. et al. RecA stimulates sister chromatid exchange and the fidelity of double-strand break repair, but not gene targeting, in plants transformed by agrobacterium. Proc Natl Acad Sci USA 97:3358-3363, 2000.

65) Richmond, E. & Peterson, C.L. Functional analysis of the DNA-stimulated ATPase domain of yeast SWI2/SNF2. Nucl Acids Res 24:3685-92, 1996.

66) Schaffer, R. et al. The late elongated hypocotyl mutation of Arabidopsis disrupts circadian rhythms and the photoperiodic control of flowering. Cell 93:1219-29, 1998.

67) Shalev, G. et al. Stimulation of homologous recombination in plants by expression of

the bacterial resolvase ruvC. Proc Natl Acad Sci USA 96:7398-7402, 1999.

68) Shen, X. et al. A chromatin remodeling complex involved in transcription and DNA processing. Nature 406:541-544, 2000.

69) Shen, X. et al. Modulation of ATP-dependent chromatin-remodeling complexes by inositol polyphosphates. Science 299: 112-114, (2003)

70) Simon, M. et al. The 3' to 5' exonuclease activity located in the DNA polymerase δ subunit of Saccharomyces cerevisiae is required for accurate replication. EMBO J 10: 2165-2170, 1991.

71) Sitney, K.C. et al. DNA polymerase III, a second essential DNA polymerase, is encoded by the S. cerevisiae CDC2 gene. Cell 56:599-605, 1989.

72) Steger, D.J. et al. Regulation of chromatin remodeling by inositol polyphosphates. Science 299: 114-116, 2003.

73) Sugino, A. Yeast DNA polymerases and their role at the replication fork. TIBS 20:319-323, 1995.

74) Swoboda, P. et al. Intrachromosomal homologous recombination in whole plants. EMBO J 13:484-489, 1994.

75) *Thykjaer, T. et al. Gene targeting approaches using positive-negative selection and* large flanking regions. Plant Mol Biol 35:523-530, 1997.

76) Torres-Ramos, C.A. et al. Requirement of Yeast DNA polymerase δ in post-replicational repair of UV-damaged DNA. J Biol Chem 272:25445-25448, 1997.

77) Travers, A. An engine for nucleosome remodeling. Cell 96:311-314, 1999.

78) Troelstra, C. et al. ERCC6, a member of a subfamily of putative helicases, is involved in Cockayne's syndrome and preferential repair of active genes. Cell 71:939-53, 1992.

79) van Gool, A.J. et al. RAD26, the functional S. cerevisiae homolog of the Cockayne syndrome B gene ERCC6. EMBO J 13:5361-5369, 1994.

80) *Walden, R. et al. Activation tagging: a means of isolating genes implicated as playing a* role in plant growth and development. Plant Mol Biol 26:1521-1528, 1994.

81) Weigel, D. et al. Activation tagging in Arabidopsis. Plant Physiol 122:1003-1013, 2000.

82) Wilson, K. et al. A Dissociation insertion causes a semidominant mutation that increases expression of TINY, an Arabidopsis gene related to APETALA2. Plant Cell 8: 659-671, 1996.

83) Wood, M. et al. An ATPase/helicase complex is an essential cofactor for oncogenic transformation by c-Myc. Mol Cell 5:321-330, 2000.

84) Zhu, T. et al. Targeted manipulation of maize genes in vivo using chimeric RNA/DNA

oligonucleotides. Proc Natl Acad Sci USA 96:8768-8773, 1999.

85) Zhu, T. et al. Engineering herbicide-resistant maize using chimeric RNA/DNA oligonucleotides. Nat Biotechnol 18:555-558,2000.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts an alignment of an AtIno80 sequence (SEQ ID NO:1) and public sequence, *At3g57300 (SEQ ID NO:3), showing a splicing difference ("Query" refers to AtIno80* sequence; "Sbjct" to public database sequence, gi|18410689|ref|NM_115590.1| (AGI:At3g57300)

SUMMARY OF THE INVENTION

The present invention provides an isolated nucleic acid, in particular DNA, comprising a *sequence having 98.5% or more identity with the sequence depicted in SEQ ID NO: 1* (AtIno80). Also provided are vectors and host cells comprising the nucleic acids of the invention, as well as polypeptides encoded by the nucleic acids.

In a further aspect of the invention, a method for inducing homologous recombination in a cell is provided, comprising modulating the expression or properties of one or more gene products selected from the group consisting of AtIno80 (SEQ ID NOs:1 and 2), At3g57300 *(SEQ ID NO:3), Rvb1 (At5g22330; SEQ ID Nos: 4 and 5), Rvb21 (At5g67630; SEQ ID NOs:* 6 and 7), Rvb22 (At3g49830: SEQ ID NOs: 8 and 9), At3g57290 (SEQ ID NO: 10), AtArp5.1 (At3g12380; SEQ ID NOs: 11 and 12), AtArp5.2 (At5g56180; SEQ ID NOs: 13 and 14), AtArp5.3 (At3g60830; SEQ ID Ns: 15 and 16) and AtArp8 (At5g43500; SEQ ID Nos: 17 and 18), their homologues, fragments or derivatives. In one embodiment, modulation is achieved by increasing expression of the gene product, such as by introducing a nucleic acid encoding the gene product into the cell operably linked to a promoter; and allowing transcription and translation of the gene in an amount sufficient to affect homologous recombination in said cell.

The method can be used to increase somatic homologous recombination and/or meiotic homologous recombination. The promoter can be an inducible promoter, a tissue-specific promoter, a constitutive promoter or a meiosis-specific promoter, depending on the desired effect.

Also provided is a method of increasing gene targeting to a desired locus in a host cell comprising introducing a desired gene into a host cell, modulating the expression or properties of one or more gene products selected from the group consisting of AtIno80, At3g57300, Rvb1, Rvb21, Rvb22, At3g57290, AtArp5.1, AtArp5.2, AtArp5.3 and AtArp8, or functional fragments, derivatives and homologues thereof in the host cell, and detecting integration of the desired gene at a selected locus in the genome of the host cell.

DETAILED DESCRIPTION OF THE INVENTION

The present inventors have used a direct screening approach to identify mutants of *Arabidopsis thaliana* showing increased frequencies of somatic recombination, by visualizing recombination events in living plants from a mutagenized population and directly isolating plants with the desired phenotype. The description below describes a genetic screen and an Arabidopsis mutant *sm22* derived from it, and the associated plant genes responsible for the altered recombination phenotype.

Existing technologies for gene targeting in plants are very inefficient. The modulation of the expression or properties of one or more gene products selected from the group consisting of AtIno80 (SEQ ID NOs:1 and 2), At3g57300 (SEQ ID NO:3), Rvb1 (At5g22330; SEQ ID Nos: 4 and 5), Rvb2(1 and 2; also referred to herein as Rvb21, At5g67630; SEQ ID NOs: 6 and 7 and Rvb22, At3g49830: SEQ ID NOs: 8 and 9, respectively), At3g57290 (SEQ ID NO: 10), AtArp5.1 (At3g12380; SEQ ID NOs: 11 and 12), AtArp5.2 (At5g56180; SEQ ID NOs: 13 and 14), AtArp5.3 (At3g60830; SEQ ID Ns: 15 and 16) and AtArp8 (At5g43500; SEQ ID Nos: 17 and 18), their homologues (including orthologs), fragments or derivatives increases the efficiency of gene targeting events and facilitates the routine manipulation of the genome of *higher plants by homologous recombination. For the purposes of this disclosure, to avoid* repetition, reference to the above group of gene products is meant to include reference to each gene individually, i.e., the modulation of the expression or properties of AtIno80, the modulation of the expression or properties of At3g57300, and so on.

An *in vivo* screen for *Arabidopsis* mutants has been devised to allow direct detection of mutants with altered recombination. As a result of the screen, mutant plants with a more than 10-fold increased or altered frequency of somatic recombination events are provided, as well as the plant gene AtIno80. One or more of AtIno80, At3g57300, Rvb1, Rvb21,

Rvb22, At3g57290, AtArp5.1, AtArp5.2, AtArp5.3 and AtArp8 or orthologs from other plant species are affected in these mutant plants. The screen allows the identification of mutant plants, and plant genes with a strong effect on recombination having little or no undesired side effects on the plant. An increase in homologous recombination frequency is useful to achieve an increased efficiency of gene targeting in plants.

Within the context of the present invention reference to a gene is to be understood as *reference to a DNA coding sequence associated with regulatory sequences*, which allow transcription of the coding sequence into RNA such as mRNA, rRNA, tRNA, snRNA, sense RNA or antisense RNA. Examples of regulatory sequences are promoter sequences, 5' and 3' untranslated sequences, introns, and termination sequences.

A promoter is understood to be a DNA sequence initiating transcription of an associated DNA sequence, and may also include elements that act as regulators of gene expression such as activators, enhancers, or repressors.

Expression of a gene refers to its transcription into RNA or its transcription and subsequent translation into protein within a living cell. In the case of antisense constructs expression refers to the transcription of the antisense DNA only.

The term transformation of cells designates the introduction of nucleic acid into a host cell, particularly the stable integration of a DNA molecule into the genome of said cell.
Any part or piece of a specific nucleotide or amino acid sequence is referred to as a component sequence or fragment.

In one aspect of the invention, nucleic acids and polypeptides are provided that can modulate homologous recombination. A nucleic acid according to the present invention comprises a sequence having 98.5%, 99%, 99.5% or more identity with the sequences depicted in SEQ ID NO:1. The nucleic acid can be DNA or RNA, such as, mRNA, rRNA, tRNA, snRNA, sense RNA or antisense RNA. Also provided is a vector comprising the nucleic acid of the invention, as well as host cells comprising the vector or nucleic acid of the invention. Suitable vectors and host cells are described in more detail below. Also provided are polypeptides encoded by the nucleic acids of the invention.

In a further aspect of the invention, methods for increasing homologous recombination are provided by modulating the expression or properties of one or more gene products selected from the group consisting of AtIno80, At3g57300, Rvb1, Rvb21, Rvb22, At3g57290, AtArp5.1, AtArp5.2, AtArp5.3 and AtArp8. In order to increase homologous recombination several methods are useful depending on the gene and the gene targeting technique employed. Typically, modulation will mean increasing the activity of the gene product, which can easily be achieved by methods known in the art.

In one embodiment, the desired gene is overexpressed in a host cell in an amount sufficient to increase homologous recombination in the host cell. By "overexpression", it is meant increasing the amount of desired gene product in a host cell, compared to untreated cells. A simple way to achieve overexpression is to produce transgenic host cells, in particular transgenic plants, carrying a construct (vector) that ectopically overexpresses the sequence of interest under the control of a suitable promoter, such as the 35S CaMV, MAS (mannopine synthase) or ubiquitin promoter.

In another embodiment, an inducible promoter is used to allow an increase in homologous recombination frequency at the time and place needed, for example, for gene targeting.

Alternatively, the construct increasing recombination can be provided at the same time as the targeting construct by co-transformation, the effect is then achieved by the transient expression of the construct containing the said genes.

Functional fragments, homologues (including orthologs) or derivatives can be easily identified by alignment with the sequences referred to above. In general two approaches exist to sequence alignment. Algorithms as proposed by Needleman & Wunsch and by Sellers align the entire length of two sequences providing a global alignment of the sequences. The Smith-Waterman algorithm on the other hand yields local alignments. A local alignment aligns the pair of regions within the sequences that are most similar given the choice of scoring matrix and gap penalties. This allows a database search to focus on the most highly conserved regions of the sequences. It also allows similar domains within sequences to be identified. To speed up alignments using the Smith-Waterman algorithm both BLAST (Basic Local Alignment Search Tool) and FASTA place additional restrictions on the alignments.

Within the context of the present invention alignments are conveniently performed using BLAST, a set of similarity search programs designed to explore all of the available sequence databases regardless of whether the query is protein or DNA. Version BLAST 2.0 (Gapped BLAST) of this search tool has been made publicly available on the internet (currently http://www.ncbi.nlm.nih.gov/BLAST/). It uses a heuristic algorithm which seeks local as opposed to global alignments and is therefore able to detect relationships among sequences which share only isolated regions. The scores assigned in a BLAST search have a well-defined statistical interpretation. Particularly useful within the scope of the present invention are the blastp program allowing for the introduction of gaps in the local sequence alignments and the PSI-BLAST program, both programs comparing an amino acid query sequence against a protein sequence database, as well as a blastp variant program allowing local alignment of two sequences only. Said programs are preferably run with optional parameters set to the default values.

Sequence alignments using BLAST can also take into account whether the substitution of one amino acid for another is likely to conserve the physical and chemical properties necessary to maintain the structure and function of the protein or is more likely to disrupt essential structural and functional features of a protein. Such sequence similarity is quantified in terms of a percentage of "positive" amino acids, as compared to the percentage of identical amino acids and can help assigning a protein to the correct protein family in border-line cases.

Specific examples of DNA and encoded proteins according to the present invention are described in SEQ ID NOS: 1-18. The putative ATPase/helicase Atlno80 may, as demonstrated in yeast, be part of a complex containing one or more of Rvb1, Rvb2, Arp5 and Arp8. All these proteins may be useful in increasing homologous recombination frequency.

Typically, functional fragments or derivatives are characterized by an amino acid sequence comprising a component sequence of at least 150 amino acid residues having 40% or more identity with an aligned component sequence of the one or more of the polypeptides encoded by the DNA of SEQ ID NOs: 1, 3, 4, 6, 8, 10, 11, 13, 15, 16 or 18. Preferably the amino acid sequence identity is higher than 50% or even higher than 55%. Most preferably

the protein sequence is that of SEQ ID NO:2.

DNA encoding proteins according to the present invention can be isolated from monocotyledonous and dicotyledonous plants. Preferred sources are corn, sugarbeet, sunflower, winter oilseed rape, soybean, cotton, wheat, rice, potato, broccoli, cauliflower, cabbage, cucumber, sweet corn, daikon, garden beans, lettuce, melon, pepper, squash, tomato, or watermelon. However, they can also be isolated from mammalian sources such as mouse or human tissues. The following general method, can be used, which the person skilled in the art knows to adapt to the specific task. A single stranded fragment of the desired gene consisting of at least 15, preferably 20 to 30 or even more than 100 consecutive nucleotides is used as a probe to screen a DNA library for clones hybridizing to said fragment. The factors to be observed for hybridization are described in Sambrook et al, Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, chapters 9.47-9.57 and 11.45-11.49, 1989. Hybridizing clones are sequenced and DNA of clones comprising a complete coding region encoding a protein characterized by an amino acid sequence comprising a component sequence of at least 150 amino acid residues having 40% or more sequence identity to the protein sequence encoded by the desired gene is purified. Said DNA can then be further processed by a number of routine recombinant DNA techniques such as restriction enzyme digestion, ligation, or polymerase chain reaction analysis. The disclosure of the nucleotide sequences in SEQ ID NOs: 1, 3, 4, 6, 8, 10, 11, 13, 15, 16 and 18 enables a person skilled in the art to design oligonucleotides for polymerase chain reactions which attempt to amplify DNA fragments from templates comprising a sequence of nucleotides characterized by any continuous sequence of 15 and preferably 20 to 30 or more basepairs of the desired gene.

Suitable vectors for practicing the methods of the invention are well known in the art. Similalry, host cells can be derived from monocotyledonous or dicotyledonous plants. Preferred sources are corn, sugarbeet, sunflower, winter oilseed rape, soybean, cotton, wheat, rice, potato, broccoli, cauliflower, cabbage, cucumber, sweet corn, daikon, garden beans, lettuce, melon, pepper, squash, tomato, or watermelon. However, host cells can also be isolated from other sources, including mammalian sources such as mouse or human cells, in particular stem cells. It is preferred that mammalian homologues are used in mammalian cells.

The methods for increasing homologous recombination are useful to obtain gene targeting so that a gene of interest is introduce into the genome at a desired locus, instead of randomly. For some hosts, in particular crop plants, the gene is preferably expressed in a selected tissue where expression is needed. This is easily achieved by the use of tissue specific promoter. Thus, the present invention provides a method for increasing somatic homologous recombination and increasing gene targeting by modulating the expression or properties of one or more gene products selected from the group consisting of AtIno80, At3g57300, Rvb1, Rvb21, Rvb22, At3g57290, AtArp5.1, AtArp5.2, AtArp5.3 and AtArp8 and fragments, derivatives and homologues thereof, essentially as described above. As is apparent to one of ordinary skill in the art, the corresponding ortholog is preferably used for any particular plant. For example, the corn ortholog of Ino80 is used (or modulated) to increase recombination in corn.

The methods are also useful to improve meiotic recombination, thereby facilitating breeding of species, in which genes encoding a particular phenotype are transferred between plants. Crossing in an interesting trait from another variety or species into a given variety by conventional breeding is a very time and labour-intensive process. Several generations of back-crosses have to be carried out to eliminate the undesired genetic material of the donor species, while maintaining the desired phenotype or trait. Using the methods described above for increasing homologous recombination, meiotic recombination frequencies can be increased, preferably by expressing the desired gene under the control of a meiosis-specific promoter or inducible promoter, the breeding process is speeded up. Thus, the present invention provides a method for increasing meiotic recombination by modulating the expression or properties of one or more gene products selected from the group consisting of AtIno80, At3g57300, Rvb1, Rvb21, Rvb22, At3g57290, AtArp5.1, AtArp5.2, AtArp5.3 and AtArp8 and fragments, derivatives and homologues thereof, essentially as described above.

The Examples below are provided for illustrative purposes and are in no way intended to be limiting to the invention.

EXAMPLES:

Example 1: Identification of sm22 gene effective in increasing homologous gene recombination.

We have used for our screening a newly constructed a transgenic Arabidopsis thaliana line that carries a recombination reporter construct based on the firefly luciferase gene. The structure of the reporter construct - two segments of the luciferase gene arranged as inverted repeats - is comparable to that of the previously described beta-glucuronidase reporter (Swoboda et al. 1994, Puchta et al. 1995a, Puchta et al. 1995b), but offers the advantage that recombination events can be detected in living plants. Luciferase activity in cells in which recombination has restored an intact luciferase gene can be detected by light emission after application of the substrate D-luciferin using a high-sensitivity CCD camera (Millar et al. 1992, Millar et al. 1995a, Millar et al. 1995b, Michelet & Chua 1996).

To induce hyperrecombination mutations in the luciferase recombination reporter line, we used T-DNA activation tagging with a mutagenic construct (pAC102). "Activation tagging" refers to the transcriptional activation of endogenous plant genes by random integration of a construct that carries promoter or enhancer sequences. One published approach for "activation tagging" is the introduction, via Agrobacterium-mediated gene transfer, of a T-DNA carrying several copies of the cauliflower mosaic virus (CaMV) 35S enhancer (Fang et al. 1989), which can activate the expression of heterologous genes over a distance (Hayashi et al. 1992, Walden et al. 1994, Kakimoto 1996, Kardailsky et al. 1999, Weigel et al. 2000). Another published approach is the introduction of a complete, outward-pointing CaMV 35S promoter on a transposable Ds element (Wilson et al. 1996, Schaffer et al. 1998, Fridborg et al. 1999). The construct "pAC102" used for our experiments is a combination of these previously described elements: it is a binary vector carrying a T-DNA that can be transferred to plants that contains a complete, outward-pointing copy of the CaMV 35S promoter/enhancer close to the right T-DNA border. Thus, this construct combines the ease of application of T-DNA gene transfer with the genetic ability of a complete promoter, avoiding some of the drawbacks of enhancer-only constructs (Weigel et al. 2000).

In principle, the activation tagging construct can cause several kinds of mutations after integration in the plant genome: gene disruption by insertion within a coding sequence,

activation of plant gene expression by action of the CaMV 35S enhancer, direct expression of a plant gene from the CaMV 35S promoter on the T-DNA, or down-regulation of expression by antisense RNA production driven from the CaMV 35S promoter. The pAC102 T-DNA carries in addition to the 35S promoter a complete copy of the pUC cloning vector to facilitate gene cloning by plasmid rescue (Dilkes & Feldmann 1998), and a sulfonamide resistance marker (Guerineau et al. 1990, Reiss et al. 1996) for selection of transgenic plants.

We transformed 13.000 three-week old Arabidopsis ecotype Columbia plants from the luciferase recombination reporter line with the activation-tagging T-DNA construct "pAC102" by Agrobacterium-mediated gene-transfer, using the established "floral dip" procedure (Clough & Bent 1998) with a modified infiltration buffer, in which the Silwet L-77 detergent was replaced by 0.05% Extravon® (Ciba). Seeds from the infiltrated plants were harvested three weeks after infiltration. Transgenic progeny carrying the pAC102 activation tagging T-DNA were selected by sowing seeds on perlite substrate drenched with Gamborg B5 mineral medium (Gamborg et al. 1968) containing 10 mg/l sulfadiazine (Sigma), and transferring surviving individuals after 10 days to soil. About 20,000 sulfonamide-resistant plants were isolated; they represent independent transformants with the pAC102 T-DNA activation tagging construct integrated at different random positions in the Arabidopsis genome.

When individual plants had grown to the 10-leaf stage, they were assayed for luciferase activity to detect somatic recombination events. Batches of 25 plants were sprayed with the substrate D-luciferin and pictures (typically two) were taken with a "Astrocam" (Gloor Instruments, Uster) by integrating photons over 15 min. Background noise and cosmic radiation was filtered out by correlating both images using the minimum function. Plants showing an increased number of sectors with luciferase activity relative to the average of the population were observed with a frequency of about 1 in 500 plants.

As a result of the screen, a hyperrecombination mutant plant was isolated called *sm22*. The original hyper-recombination phenotype of sm22 plant shows an enhancement of about 20- to 50- fold for homologous recombination in the reporter line. No other obvious phenotype was seen and the seed yield was normal. Sulfonamide selection in the second generation (T2) revealed a 2/1 or 3/1 segregation of resistant seedlings, thus showing that there is only 1 locus (or 2 closely related loci) with an active T-DNA inserted. However, the T2

recombination phenotype was even lower (less or same number of recombination events per plant) than in the wild type.

After HindIII digestion of T1 callus genomic DNA prepared essentially according to the method of the Nucleon Phytopure protocol and Plant DNA extraction kit (Amersham), plasmid rescue was applied (Dilkes & Feldmann 1998, Mathur et al. 1998), which gave rise to two independent junction fragments. Briefly, we digested plant genomic DNA with HinDIII, circularized the resulting fragments by ligation at low DNA concentration, and transformed the ligation mixture into competent *E. coli* TOP10 cells (commercially available from INVITROGEN) by electroporation. Since the HindIII fragments containing the fusion joint between plant DNA and the right end of the activation tagging construct carry a plasmid origin and the ampicillin resistance gene (bla) contributed by pAC102, circularization of such fragments will result in a functional bacterial plasmid and confer ampicillin resistance to the *E. coli* cells.

Several colonies were obtained after plating the transformed bacteria on selection medium containing ampicillin. Plasmid DNA of these transformants was prepared and characterized by restriction analysis.To determine the nature of the plant sequences joined to the right end of the T-DNA, the plant DNA insert from these rescued plasmids was sequenced from both sides, using one custom sequencing primer complementary to the T-DNA right end reading towards the plant DNA, and the standard M13 reverse sequencing primer, reading from the pAC102 vector sequences into the plant DNA insert from the other end. The obtained DNA sequences were compared to the GenBank nucleotide database using the BLASTN search program.

Two insertions were identified. The first one corresponds to a single T-DNA insertion without deletion (left border, LB, junction sequenced) in the N-terminal region of a putative ATPase/helicase gene, in antisense orientation. The second T-DNA inserted in a gene with no obvious relationship to homologous recombination (gb AF082176_1) and does not confer sulfonamide resistance. Six (T3) resistant families were analysed by PCR and Southern. Only one family contained some plants with the second insertion whereas all families have the helicase insertion site.

In subsequent generations, homozygous plants for the helicase insertion site were obtained. The homologous recombination frequency of heterozygous and homozygous plants for this insertion site was at least 50% and 15%, respectively, and up to 80% and 20%, respectively, of the wild type level.

The complete cDNA (4.8kb) was cloned in two steps. First, a public EST containing the 3' part was sequenced. Then the 5' part of the cDNA was amplified by RT-PCR on Col-0 (*Arabidopsis* Columbia ecotype, wild type) callus RNA (prepared with the Qiagen RNAeasy Plant Kit), using primers in the 5' untranslated region including a stop codon in frame with the predicted ATG (sm5UT) to make sure that the complete 5' part of the cDNA was amplified. The primer sequences were sm5UT: ctagaagcttttaaggat*TAA*gactctcc (SEQ ID NO:19) and for 3' primer: ctcgtatgtatccccttctcc (SEQ ID NO:20). The coding sequence is provided as SEQ ID NO.1 and is similar to but not identical to At3g57300 (see Fig.1).

The predicted helicase gene (8kb genomic DNA) has about 23 exons encoding a protein of 1507 amino acids (SEQ ID NO:2). It is predicted to be an ATPase of the Swi2/Snf2 family, and contains several nuclear localization signals (NLS). The ATPase/helicase gene is the putative Arabidopsis ortholog of the yeast Ino80p/YGL150c protein (Ebbert et al. (1999), Shen et al. (2000)). Homologs exist in yeast, budding yeast, *Drosophila* and human. These four homologues have several highly conserved regions including the six motifs of the SWI2/SNF2 helicase domain. Several NLS suggest a nuclear localization of the gene product.

In the sm22 heterozygous and homozygous mutant, the level of Ino80 transcript is respectively about 50% of the wild type situation or absent, as measured by semi-quantitative RT-PCR (5' At3g57300 primer: TGATGGATCTATCACCATCAG, SEQ ID NO:21; 3' At3g57300 primer: ggtgggattccaatcactttc, SEQ ID NO:22) and by northern blot hybridization. For this, the RNA was extracted from 2 weeks old seedlings using the QIAGEN RNAeasy plant extraction kit following the manufacturer's instructions. Together with the decrease of homologous recombination in sm22 plants described above (50% in heterozygous plants, 15% in homozygous mutant), this result shows that the level of Ino80 gene product might positively and directly regulate the homologous recombination frequency, making this gene a choice candidate to positively regulate homologous recombination.

Results with a recombination reporter line 1445 (Gherbi et al. 2001) overexpressing the INO80 cDNA under the control of the 35S promoter and with an N-terminal HA-tag interrupted by an intron show upregulation of homologous recombination providing evidence that INO80 upregulates homologous recombination.

The yeast homolog (Ebbert *et al.*,1999), INO80(=YGL150c) is part of a big complex >1MDa (monomeric form is 171KDa), containing two essential helicases Rvb1p and Rvb2p and actin related proteins (arp) Arp4, Arp5 and Arp8 (Cho et al. 2001; Jonsson et al. 2001; Wood et al. 2000). Thus, the involvement of Ino80 in homologous recombination implicates the activity of these other genes in homologous recombination. In eukaryotes Human Rvb1p and Rvb2p are also known (Kanemaki 1999, Ikura et al. 2000, Shen et al. 2000).

In *Arabidopsis thaliana* we found three genes closely related to Rvbs from other organisms, AtRvb1 (SEQ ID NO:4, SEQ ID NO:5), AtRvb21 (SEQ ID NO: 6, SEQ ID NO:7) and AtRvb22 (SEQ ID NO:8, SEQ ID NO:9). The three genes are expressed (RT-PCR) and some of them are positively regulated by genotoxic stress (UVc, bleomycin). For treatment with Bleomycin (BLM) 2 week-old Arabidopsis seedlings were placed under sterile conditions in liquid GM medium containing 10^{-6}M of BLM (Sigma) or 100 ppm of MMS (Fluka, Switzerland). For UV-C irradiation (6000 ergs) 2 week-old seedlings were irradiated with light provided by a HNS 55W OFR lamp (Osram). After treatment, plants were harvested at several time points (30min, 1h, 4h and 12h) and RNA extracted as described above. Then semi-quantitative RT-PCR analysis was performed with the following primers AtIno80 (TGATGGATCTATCACCATCAG, SEQ ID NO:23; ggtgggattccaatcactttc, SEQ ID NO:24) AtRvb1 (tttgatgggccaaatgatg, SEQ ID NO:25; cttccaaCCTAGGtgagatgtttcaacaaaatgtgc, SEQ ID NO:26) AtRvb21 (tcaacagcaggacacaagg, SEQ ID NO:27; cccaatgCCTAGGaaatccgagttcaacatcctaatc, SEQ ID NO:28) AtRvb22 (acaaaccagatatcagcacatgg, SEQ ID NO:29; aacaagtactcgctctcatgctc, SEQ ID NO:30).

To characterize further the *Atino80*-1 HR deficiency, we subjected the mutant to various genotoxic stresses. In parallel with the original *ino80* mutant we also tested two allelic mutants of *INO80* from the publicly available SAIL mutant collection. Neither bleomycin nor Mitomycin-C or UV-C sensitivity was shifted in the *Atino80*-1 mutants, in any of the various conditions tested. All the *Atino80*-1 alleles seem to be slightly hypersensitive to MMS (methyl

methanesulfonate), which is also known to induce DNA double-strand breaks. The difference in sensitivity is visible at 60 and 80ppm of MMS on root elongation and rosette growth. Most mutations affecting DNA repair or recombination also give rise to changes in the sensitivity to DNA damaging agents. We challenged *Atino80*-1 mutant plants with various treatments known to induce DNA damage and recombination. None of the tested agents (UV-C, bleomycin, Mitomycin-C and MMS) gave rise to a shift in sensitivity, with the exception of a slight change for MMS. This suggests a difference for the role of INO80 in plants compared to yeast (Ebbert et al., 1999; Shen et al. 2000) and supports the use of AtINO80 to regulate homologous recombination without affecting the major repair pathway in plants.

In the sm22 background the steady state level of AtRvb21 and AtRvb22 was shown to be down-regulated using RT-PCR on RNA extracted as above mentioned.
This indicates that the components of the putative Arabidopsis INO80 complex show co-regulation at the transcriptional level, supporting the use of Arabidopsis Rvb1, Rvb21 and Rvb22 and the *Arabidopsis* Arp protein orthologs to manipulate homologous recombination frequency in plants.

Example 2: AtRvb1 as positive regulator of homologous recombination.

As describe above (Example 1), the original recombination-up phenotype found in sm22 can be associated with an effect mediated by the Arabidopsis Rvb1 and 2 orthologs. Thus, *AtRvb1 can be used as a positive regulator of homologous recombination.*

Example 3: AtRvb21 as positive regulator of homologous recombination.

As describe above (Example 1), the original recombination-up phenotype found in sm22 can be associated with an effect mediated by the Arabidopsis Rvb1 and 2 orthologs. Thus, AtRvb21 can be used as a positive regulator of homologous recombination.

Example 4: AtRvb22 as positive regulator of homologous recombination.

As describe above (Example 1), the original recombination-up phenotype found in sm22 can be associated with an effect mediated by the Arabidopsis Rvb1 and 2 orthologs. Thus, AtRvb22 can be used as a positive regulator of homologous recombination.

Example 5: **At3g57290 as positive regulator of homologous recombination.**

In the *sm22* mutant (Example 5), the At3g57290p gene is potentially overexpressed by the 35S Enhancer/promoter. Over expression of this gene in the sm22 context or directly with a 35S promoter can be carried out to reproduce the original recombination-up phenotype. The phenotype was lost in the second generation (Example 1), at which point At3g57290 is not overexpressed any longer allowing a temporal ability to modulate homologous recombination.

Example 6: AtArp as positive regulators of homologous recombination.

As describe above (Example 1), the original recombination-up phenotype found in sm22 can be associated with an effect mediated by other components of the Arabidopsis INO80 complex, such as the Arp homolog AtArp5.1, AtArp5.2, AtArp5.3 and/or AtArp8. Any of these Arp hmologues can be used as a positive regulator of homologous recombination, alone or in combination.

All publications referred to herein as well as the disclosure of GB patent application 0214896.3 are incorporated by reference as if each is referred to individually.

What is claimed is:

1. An isolated nucleic acid comprising a sequence having 98.5% or more identity with the sequence depicted in SEQ ID NO:1.

2. The nucleic acid of claim 1, wherein said nucleic acid is DNA.

3. A vector comprising the nucleic acid if claim 2.

4. A host cell comprising the vector or nucleic acid of claim 3.

5. A polypeptide encoded by the isolated nucleic acid of claim 1.

6. A method for inducing homologous recombination in a cell, said method comprising modulating the expression or properties of one or more gene products selected from the group consisting of AtIno80, At3g57300, Rvb1, Rvb21, Rvb22, Arp5 and Arp8, fragments, derivatives and homologues thereof,.

7. The method of claim 11, said method comprising increasing expression of said gene product.

8. The method of claim 12, said method comprising introducing a nucleic acid encoding said gene product into said cell operably linked to a promoter; and allowing transcription and translation of said gene in an amount sufficient to affect homologous recombination in said cell.

9. The method of claim 13, wherein said homologous recombination is somatic homologous recombination.

10. The method of claim 13, wherein said homologous recombination is meiotic homologous recombination.

11. The method of claim 13, wherein said promoter is an inducible promoter.

12. The method of claim 13, wherein said promoter is a tissue-specific promoter.

13. The method of claim 13, wherein said promoter is a constitutive promoter.

14. The method of claim 13, wherein said promoter is a meiosis-specific promoter.

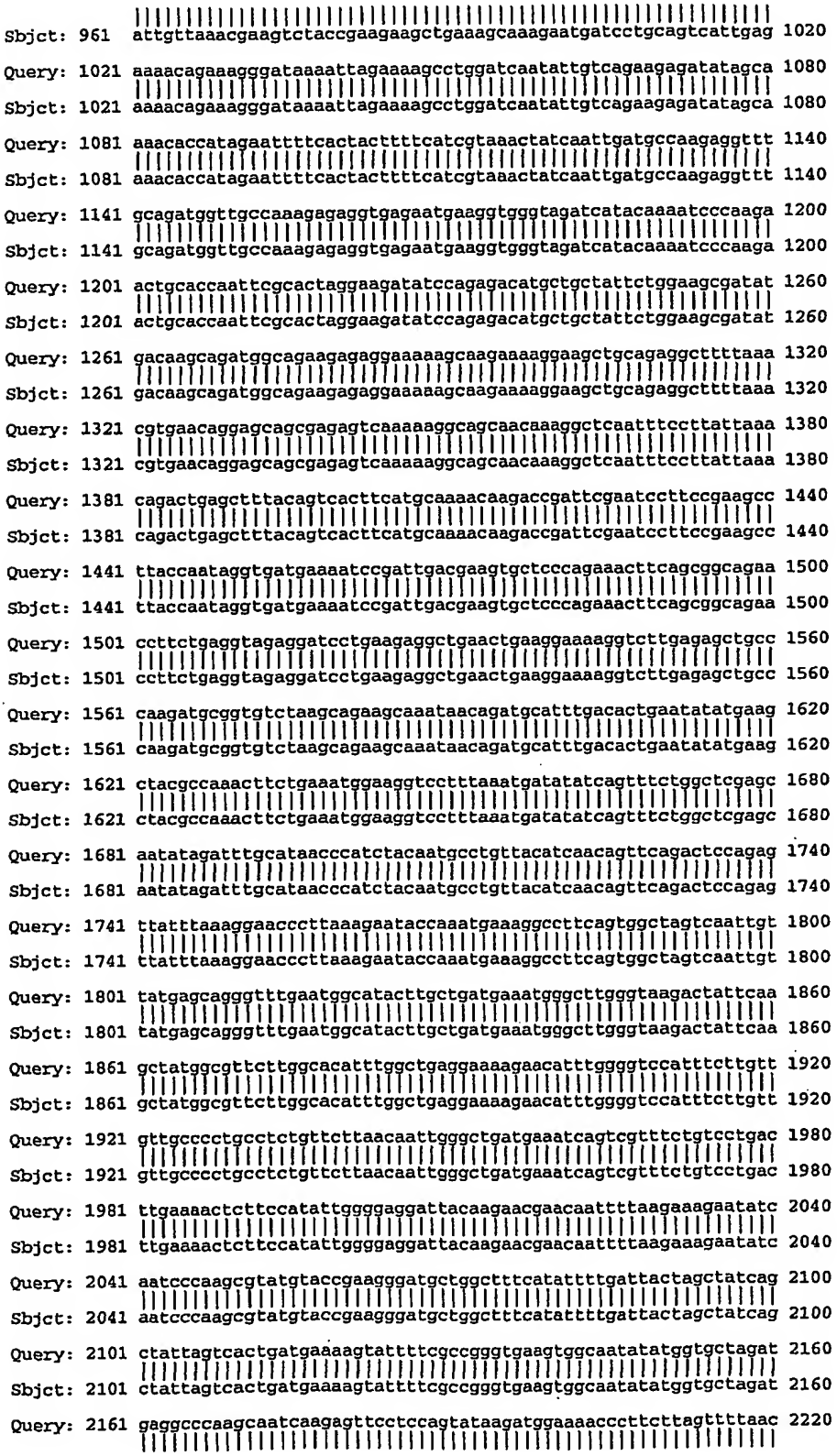
15. A method of increasing gene targeting to a desired locus in a host cell, said method comprising introducing a desired gene into a host cell, modulating the expression or properties of one or more gene products selected from the group consisting of AtIno80, At3g57300, Rvb1, Rvb21, Rvb22, Arp5 and Arp8 or functional fragments, derivatives and homologues thereof in said host cell, and detecting integration of said desired gene at a selected locus in the genome of said host cell.

Figure 1

>Alignment of AtIno 80 sequence and public sequence, At3g57300, showing splicing difference

Query: claimed sequence
Sbjct: gi|18410689|ref|NM_115590.1| (AGI:At3g57300)

```
Query: 1    atggatccttcaagacgaccaccgaaggactctccttacgcgaatctattcgatctcgag 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 1    atggatccttcaagacgaccaccgaaggactctccttacgcgaatctattcgatctcgag 60

Query: 61   ccgttgatgaagtttagaattccgaaacctgaagatgaagttgattattatgggagtagt 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 61   ccgttgatgaagtttagaattccgaaacctgaagatgaagttgattattatgggagtagt 120

Query: 121  agccaggatgaaagtagaagcactcaaggtggggtagtggcaaactacagcaatgggtct 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 121  agccaggatgaaagtagaagcactcaaggtggggtagtggcaaactacagcaatgggtct 180

Query: 181  aaatcgagaatgaatgcgagctccaagaagagaaagcggtggacagaagctgaggatgca 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 181  aaatcgagaatgaatgcgagctccaagaagagaaagcggtggacagaagctgaggatgca 240

Query: 241  gaggacgatgatgatctctacaatcaacatgttactgaggagcactaccgatcaatgctt 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 241  gaggacgatgatgatctctacaatcaacatgttactgaggagcactaccgatcaatgctt 300

Query: 301  ggggagcatgtacaaaaattcaaaaataggtccaaggagactcaagggaatcctcctcat 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 301  ggggagcatgtacaaaaattcaaaaataggtccaaggagactcaagggaatcctcctcat 360

Query: 361  ctgatgggttttccggtgctaaagagcaatgtgggcagttacagaggtaggaaaccaggg 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 361  ctgatgggttttccggtgctaaagagcaatgtgggcagttacagaggtaggaaaccaggg 420

Query: 421  aatgattaccatgggaggttctatgacatggacaactctccaaattttgcagctgatgtg 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 421  aatgattaccatgggaggttctatgacatggacaactctccaaattttgcagctgatgtg 480

Query: 481  accccacataggcgaggaagctaccatgatcgtgatattacacccaagatagcatatgaa 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 481  accccacataggcgaggaagctaccatgatcgtgatattacacccaagatagcatatgaa 540

Query: 541  ccttcgtatttggacattggtgatggtgtcatctacaaaatccccccaagttatgacaag 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 541  ccttcgtatttggacattggtgatggtgtcatctacaaaatccccccaagttatgacaag 600

Query: 601  ctggtggcatcattaaacttaccgagcttttcagacattcatgtggaagaattttacttg 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 601  ctggtggcatcattaaacttaccgagcttttcagacattcatgtggaagaattttacttg 660

Query: 661  aaaggaactctggatctgagatcattagcagaactgatggcaagtgataaaaggtctgga 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 661  aaaggaactctggatctgagatcattagcagaactgatggcaagtgataaaaggtctgga 720

Query: 721  gtaagaagccgtaatggaatgggtgagcctcgacctcaatatgaatctcttcaagctaga 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 721  gtaagaagccgtaatggaatgggtgagcctcgacctcaatatgaatctcttcaagctaga 780

Query: 781  atgaaggccctgtcaccttcaaactccaccccaaattttagcctcaaggtgtcagaagct 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 781  atgaaggccctgtcaccttcaaactccaccccaaattttagcctcaaggtgtcagaagct 840

Query: 841  gcaatgaattctgccattccagaaggatctgctggaagtactgcacggacaattctgtct 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 841  gcaatgaattctgccattccagaaggatctgctggaagtactgcacggacaattctgtct 900

Query: 901  gagggtggtgttttacaggtccattacgtgaagattctggagaaggggatacatacgag 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 901  gagggtggtgttttacaggtccattacgtgaagattctggagaaggggatacatacgag 960

Query: 961  attgttaaacgaagtctaccgaagaagctgaaagcaaagaatgatcctgcagtcattgag 1020
```



```
Sbjct: 2161 gaggcccaagcaatcaagagttcctccagtataagatggaaaaccccttcttagttttaac 2220

Query: 2221 tgtcggaaccgattgcttctgactggtactccaattcagaacaacatggcagagttatgg 2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2221 tgtcggaaccgattgcttctgactggtactccaattcagaacaacatggcagagttatgg 2280

Query: 2281 gccctgctgcatttcatcatgccaatgttgtttgacaaccatgatcaatttaatgaatgg 2340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2281 gccctgctgcatttcatcatgccaatgttgtttgacaaccatgatcaatttaatgaatgg 2340

Query: 2341 ttctcaaaaggaattgagaatcatgctgaacacggaggcactttaaatgagcaccagctt 2400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2341 ttctcaaaaggaattgagaatcatgctgaacacggaggcactttaaatgagcaccagctt 2400

Query: 2401 aacagactgcatgcgatcttgaaaccgttcatgcttcgacgggtaaaaaaggatgtggtt 2460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2401 aacagactgcatgcgatcttgaaaccgttcatgcttcgacgggtaaaaaaggatgtggtt 2460

Query: 2461 tctgagctaactacaaagacggaagttacagtacactgcaagctcagttctcgacaacaa 2520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2461 tctgagctaactacaaagacggaagttacagtacactgcaagctcagttctcgacaacaa 2520

Query: 2521 gctttttatcaggctattaagaacaaaatttctctggctgagttgtttgatagcaaccgc 2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2521 gctttttatcaggctattaagaacaaaatttctctggctgagttgtttgatagcaaccgc 2580

Query: 2581 ggacaatttactgataagaaagtattgaatttaatgaatattgtcattcaactaaggaag 2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2581 ggacaatttactgataagaaagtattgaatttaatgaatattgtcattcaactaaggaag 2640

Query: 2641 gtttgcaaccatccagagttgttcgaaaggaatgaagggagctcgtatctctactttgga 2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2641 gtttgcaaccatccagagttgttcgaaaggaatgaagggagctcgtatctctactttgga 2700

Query: 2701 gtgacttccaattctcttttgccccatccctttggtgagctagaggatgtacattattct 2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2701 gtgacttccaattctcttttgccccatccctttggtgagctagaggatgtacattattct 2760

Query: 2761 ggtggtcaaaatccgataatatacaagatacctaagctactacaccaagaggtgctccaa 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2761 ggtggtcaaaatccgataatatacaagatacctaagctactacaccaagaggtgctccaa 2820

Query: 2821 aattctgaaacattttgttcttctgtcgggcgtggcatctcaagagaatcttttctgaag 2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2821 aattctgaaacattttgttcttctgtcgggcgtggcatctcaagagaatcttttctgaag 2880

Query: 2881 cattttaatatatattcacctgagtatattcttaagtcaatattcccatctgatagtggg 2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2881 cattttaatatatattcacctgagtatattcttaagtcaatattcccatctgatagtggg 2940

Query: 2941 gtagatcaagtggttagtggaagtggagcatttggcttttcacgcttgatggatctatca 3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2941 gtagatcaagtggttagtggaagtggagcatttggcttttcacgcttgatggatctatca 3000

Query: 3001 ccatcagaagttggatatctggctctgtgttctgttgcagaaaggctattattttctata 3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 3001 ccatcagaagttggatatctggctctgtgttctgttgcagaaaggctattattttctata 3060

Query: 3061 ctgaggtgggagcggcaatttttggatgaattagttaactctcttatggagtccaaggat 3120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 3061 ctgaggtgggagcggcaatttttggatgaattagttaactctcttatggagtccaaggat 3120

Query: 3121 ggtgatcttagtgacaataacatcgagagagttaaaaccaaagctgtcacaagaatgttg 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 3121 ggtgatcttagtgacaataacatcgagagagttaaaaccaaagctgtcacaagaatgttg 3180

Query: 3181 ctgatgccatcaaaagttgaaacgaattttcagaaaaggagactaagcacagggcctacc 3240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 3181 ctgatgccatcaaaagttgaaacgaattttcagaaaaggagactaagcacagggcctacc 3240

Query: 3241 cgtccttcatttgaagcgctagtgatctctcatcaggataggtttctttcaagtatcaaa 3300
```

```
Sbjct: 3241 cgtccttcatttgaagcgctagtgatctctcatcaggataggtttctttcaagtatcaaa 3300

Query: 3301 ctcctgcattctgcatatacttatatcccaaaagccagagctccacctgtaagcattcat 3360
Sbjct: 3301 ctcctgcattctgcatatacttatatcccaaaagccagagctccacctgtaagcattcat 3360

Query: 3361 tgctcggacagaaattcggcatacagagttacagaagaattacatcaaccatggcttaag 3420
Sbjct: 3361 tgctcggacagaaattcggcatacagagttacagaagaattacatcaaccatggcttaag 3420

Query: 3421 agactattaatcggttttgcacgaacgtcagaagctaatggacccaggaagcctaacagc 3480
Sbjct: 3421 agactattaatcggttttgcacgaacgtcagaagctaatggacccaggaagcctaacagc 3480

Query: 3481 tttccacatcctttaatccaagaaattgattcagaacttccagttgtgcagcctgcgctt 3540
Sbjct: 3481 tttccacatcctttaatccaagaaattgattcagaacttccagttgtgcagcctgcgctt 3540

Query: 3541 caactgacacacagaatatttggttcttgccctccaatgcaaagttttgacccagcaaag 3600
Sbjct: 3541 caactgacacacagaatatttggttcttgccctccaatgcaaagttttgacccagcaaag 3600

Query: 3601 ttgctcacggactctgggaagctgcagacacttgatatattattgaagcggcttcgagct 3660
Sbjct: 3601 ttgctcacggactctgggaagctgcagacacttgatatattattgaagcggcttcgagct 3660

Query: 3661 ggaaatcacagggtgctcctgtttgcacaaatgacaaagatgctgaacattctcgaggat 3720
Sbjct: 3661 ggaaatcacagggtgctcctgtttgcacaaatgacaaagatgctgaacattctcgaggat 3720

Query: 3721 tatatgaactatagaaagtacaagtacctcaggcttgatggatcctccaccatcatggat 3780
Sbjct: 3721 tatatgaactatagaaagtacaagtacctcaggcttgatggatcctccaccatcatggat 3780

Query: 3781 cgccgagatatggttagggattttcagcataggagcgatatttttgtattcttgctgagc 3840
Sbjct: 3781 cgccgagatatggttagggattttcagcataggagcgatatttttgtattcttgctgagc 3840

Query: 3841 accagagctggaggacttggtatcaacttgacggctgcagacactgtcattttctatgaa 3900
Sbjct: 3841 accagagctggaggacttggtatcaacttgacggctgcagacactgtcattttctatgaa 3900

Query: 3901 agtgattggaatcccaccttggatttacaagctatggacagggctcatcgtcttggacag 3960
Sbjct: 3901 agtgattggaatcccaccttggatttacaagctatggacagggctcatcgtcttggacag 3960

Query: 3961 acaaaagatg 3970
Sbjct: 3961 acaaaagatg 3970


 Score = 1001 bits (505), Expect = 0.0
 Identities = 522/528 (98%), Gaps = 6/528 (1%)
 Strand = Plus / Plus

Query: 3997 gagacggtggaagagaaaattttgcacagggcaagtcagaaaaatacagttcaacagctt 4056
Sbjct: 3970 gagacggtggaagagaaaattttgcacagggcaagtcagaaaaatacagttcaacagctt 4029

Query: 4057 gttatgactggagggcatgttcagggtgatgattttcttggagctgcggatgtggtatct 4116
Sbjct: 4030 gttatgactggagggcatgttcagggtgatgattttcttggagctgcggatgtggtatct 4089

Query: 4117 ctgctaatggatgatgcggaggcagcacaactggagcagaaattcagagaactaccatta 4176
```

```
Sbjct: 4090 ctgctaatggatgatgcggaggcagcacaactggagcagaaattcagagaactaccatta 4149

Query: 4177 caggtaaaggacaggcagaagaaaaagacgaaacgtatcagaatagatgctgaaggagat 4236
             ||||      ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 4150 cagg------acaggcagaagaaaaagacgaaacgtatcagaatagatgctgaaggagat 4203

Query: 4237 gcaactttggaagagttagaagatgttgaccgacaggataacggacaggaacctttggaa 4296
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 4204 gcaactttggaagagttagaagatgttgaccgacaggataacggacaggaacctttggaa 4263

Query: 4297 gaaccggaaaagccaaaatccagtaataaaaagaggagagctgcttcaaatccgaaagct 4356
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 4264 gaaccggaaaagccaaaatccagtaataaaaagaggagagctgcttcaaatccgaaagct 4323

Query: 4357 agagctcctcagaaagcaaaggaagaagcaaatggtgaagatactcctcagaggacaaaa 4416
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 4324 agagctcctcagaaagcaaaggaagaagcaaatggtgaagatactcctcagaggacaaaa 4383

Query: 4417 agggtaaagagacaaacaaagagcataaacgaaagtcttgaacctgtattctctgcctct 4476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 4384 agggtaaagagacaaacaaagagcataaacgaaagtcttgaacctgtattctctgcctct 4443

Query: 4477 gtaacagaatcaaataaaggattcgatccaagtagctccgctaactaa 4524
             ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 4444 gtaacagaatcaaataaaggattcgatccaagtagctccgctaactaa 4491
```

SEQUENCE LISTING

<110> Novartis Forschungsstiftung, Zweigniederlassung Friedrich Miescher Institute for Biomedical Research

<120> Gene for increased somatic recombination

<130> 1-32546B

<160> 30

<170> PatentIn version 3.1

<210> 1

<211> 4524

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(4524)

<223>

<400> 1

```
atg gat cct tca aga cga cca ccg aag gac tct cct tac gcg aat cta       48
Met Asp Pro Ser Arg Arg Pro Pro Lys Asp Ser Pro Tyr Ala Asn Leu
1               5                  10                  15

ttc gat ctc gag ccg ttg atg aag ttt aga att ccg aaa cct gaa gat       96
Phe Asp Leu Glu Pro Leu Met Lys Phe Arg Ile Pro Lys Pro Glu Asp
            20                  25                  30

gaa gtt gat tat tat ggg agt agt agc cag gat gaa agt aga agc act      144
Glu Val Asp Tyr Tyr Gly Ser Ser Ser Gln Asp Glu Ser Arg Ser Thr
        35                  40                  45

caa ggt ggg gta gtg gca aac tac agc aat ggg tct aaa tcg aga atg      192
Gln Gly Gly Val Val Ala Asn Tyr Ser Asn Gly Ser Lys Ser Arg Met
    50                  55                  60

aat gcg agc tcc aag aag aga aag cgg tgg aca gaa gct gag gat gca      240
Asn Ala Ser Ser Lys Lys Arg Lys Arg Trp Thr Glu Ala Glu Asp Ala
65                  70                  75                  80
```

```
gag gac gat gat gat ctc tac aat caa cat gtt act gag gag cac tac      288
Glu Asp Asp Asp Asp Leu Tyr Asn Gln His Val Thr Glu Glu His Tyr
                85                  90                  95

cga tca atg ctt ggg gag cat gta caa aaa ttc aaa aat agg tcc aag      336
Arg Ser Met Leu Gly Glu His Val Gln Lys Phe Lys Asn Arg Ser Lys
            100                 105                 110

gag act caa ggg aat cct cct cat ctg atg ggt ttt ccg gtg cta aag      384
Glu Thr Gln Gly Asn Pro Pro His Leu Met Gly Phe Pro Val Leu Lys
        115                 120                 125

agc aat gtg ggc agt tac aga ggt agg aaa cca ggg aat gat tac cat      432
Ser Asn Val Gly Ser Tyr Arg Gly Arg Lys Pro Gly Asn Asp Tyr His
    130                 135                 140

ggg agg ttc tat gac atg gac aac tct cca aat ttt gca gct gat gtg      480
Gly Arg Phe Tyr Asp Met Asp Asn Ser Pro Asn Phe Ala Ala Asp Val
145                 150                 155                 160

acc cca cat agg cga gga agc tac cat gat cgt gat att aca ccc aag      528
Thr Pro His Arg Arg Gly Ser Tyr His Asp Arg Asp Ile Thr Pro Lys
                165                 170                 175

ata gca tat gaa cct tcg tat ttg gac att ggt gat ggt gtc atc tac      576
Ile Ala Tyr Glu Pro Ser Tyr Leu Asp Ile Gly Asp Gly Val Ile Tyr
            180                 185                 190

aaa atc ccc cca agt tat gac aag ctg gtg gca tca tta aac tta ccg      624
Lys Ile Pro Pro Ser Tyr Asp Lys Leu Val Ala Ser Leu Asn Leu Pro
        195                 200                 205

agc ttt tca gac att cat gtg gaa gaa ttt tac ttg aaa gga act ctg      672
Ser Phe Ser Asp Ile His Val Glu Glu Phe Tyr Leu Lys Gly Thr Leu
    210                 215                 220

gat ctg aga tca tta gca gaa ctg atg gca agt gat aaa agg tct gga      720
Asp Leu Arg Ser Leu Ala Glu Leu Met Ala Ser Asp Lys Arg Ser Gly
225                 230                 235                 240

gta aga agc cgt aat gga atg ggt gag cct cga cct caa tat gaa tct      768
Val Arg Ser Arg Asn Gly Met Gly Glu Pro Arg Pro Gln Tyr Glu Ser
                245                 250                 255

ctt caa gct aga atg aag gcc ctg tca cct tca aac tcc acc cca aat      816
Leu Gln Ala Arg Met Lys Ala Leu Ser Pro Ser Asn Ser Thr Pro Asn
            260                 265                 270

ttt agc ctc aag gtg tca gaa gct gca atg aat tct gcc att cca gaa      864
Phe Ser Leu Lys Val Ser Glu Ala Ala Met Asn Ser Ala Ile Pro Glu
        275                 280                 285

gga tct gct gga agt act gca cgg aca att ctg tct gag ggt ggt gtt      912
Gly Ser Ala Gly Ser Thr Ala Arg Thr Ile Leu Ser Glu Gly Gly Val
    290                 295                 300

tta cag gtc cat tac gtg aag att ctg gag aag ggg gat aca tac gag      960
Leu Gln Val His Tyr Val Lys Ile Leu Glu Lys Gly Asp Thr Tyr Glu
305                 310                 315                 320

att gtt aaa cga agt cta ccg aag aag ctg aaa gca aag aat gat cct     1008
Ile Val Lys Arg Ser Leu Pro Lys Lys Leu Lys Ala Lys Asn Asp Pro
                325                 330                 335

gca gtc att gag aaa aca gaa agg gat aaa att aga aaa gcc tgg atc     1056
Ala Val Ile Glu Lys Thr Glu Arg Asp Lys Ile Arg Lys Ala Trp Ile
```

```
                340                 345                 350

aat att gtc aga aga gat ata gca aaa cac cat aga att ttc act act     1104
Asn Ile Val Arg Arg Asp Ile Ala Lys His His Arg Ile Phe Thr Thr
        355                 360                 365

ttt cat cgt aaa cta tca att gat gcc aag agg ttt gca gat ggt tgc     1152
Phe His Arg Lys Leu Ser Ile Asp Ala Lys Arg Phe Ala Asp Gly Cys
    370                 375                 380

caa aga gag gtg aga atg aag gtg ggt aga tca tac aaa atc cca aga     1200
Gln Arg Glu Val Arg Met Lys Val Gly Arg Ser Tyr Lys Ile Pro Arg
385                 390                 395                 400

act gca cca att cgc act agg aag ata tcc aga gac atg ctg cta ttc     1248
Thr Ala Pro Ile Arg Thr Arg Lys Ile Ser Arg Asp Met Leu Leu Phe
                405                 410                 415

tgg aag cga tat gac aag cag atg gca gaa gag agg aaa aag caa gaa     1296
Trp Lys Arg Tyr Asp Lys Gln Met Ala Glu Glu Arg Lys Lys Gln Glu
            420                 425                 430

aag gaa gct gca gag gct ttt aaa cgt gaa cag gag cag cga gag tca     1344
Lys Glu Ala Ala Glu Ala Phe Lys Arg Glu Gln Glu Gln Arg Glu Ser
        435                 440                 445

aaa agg cag caa caa agg ctc aat ttc ctt att aaa cag act gag ctt     1392
Lys Arg Gln Gln Gln Arg Leu Asn Phe Leu Ile Lys Gln Thr Glu Leu
    450                 455                 460

tac agt cac ttc atg caa aac aag acc gat tcg aat cct tcc gaa gcc     1440
Tyr Ser His Phe Met Gln Asn Lys Thr Asp Ser Asn Pro Ser Glu Ala
465                 470                 475                 480

tta cca ata ggt gat gaa aat ccg att gac gaa gtg ctc cca gaa act     1488
Leu Pro Ile Gly Asp Glu Asn Pro Ile Asp Glu Val Leu Pro Glu Thr
                485                 490                 495

tca gcg gca gaa cct tct gag gta gag gat cct gaa gag gct gaa ctg     1536
Ser Ala Ala Glu Pro Ser Glu Val Glu Asp Pro Glu Glu Ala Glu Leu
            500                 505                 510

aag gaa aag gtc ttg aga gct gcc caa gat gcg gtg tct aag cag aag     1584
Lys Glu Lys Val Leu Arg Ala Ala Gln Asp Ala Val Ser Lys Gln Lys
        515                 520                 525

caa ata aca gat gca ttt gac act gaa tat atg aag cta cgc caa act     1632
Gln Ile Thr Asp Ala Phe Asp Thr Glu Tyr Met Lys Leu Arg Gln Thr
    530                 535                 540

tct gaa atg gaa ggt cct tta aat gat ata tca gtt tct ggc tcg agc     1680
Ser Glu Met Glu Gly Pro Leu Asn Asp Ile Ser Val Ser Gly Ser Ser
545                 550                 555                 560

aat ata gat ttg cat aac cca tct aca atg cct gtt aca tca aca gtt     1728
Asn Ile Asp Leu His Asn Pro Ser Thr Met Pro Val Thr Ser Thr Val
                565                 570                 575

cag act cca gag tta ttt aaa gga acc ctt aaa gaa tac caa atg aaa     1776
Gln Thr Pro Glu Leu Phe Lys Gly Thr Leu Lys Glu Tyr Gln Met Lys
            580                 585                 590

ggc ctt cag tgg cta gtc aat tgt tat gag cag ggt ttg aat ggc ata     1824
Gly Leu Gln Trp Leu Val Asn Cys Tyr Glu Gln Gly Leu Asn Gly Ile
        595                 600                 605
```

```
ctt gct gat gaa atg ggc ttg ggt aag act att caa gct atg gcg ttc     1872
Leu Ala Asp Glu Met Gly Leu Gly Lys Thr Ile Gln Ala Met Ala Phe
    610                 615                 620

ttg gca cat ttg gct gag gaa aag aac att tgg ggt cca ttt ctt gtt     1920
Leu Ala His Leu Ala Glu Glu Lys Asn Ile Trp Gly Pro Phe Leu Val
625                 630                 635                 640

gtt gcc cct gcc tct gtt ctt aac aat tgg gct gat gaa atc agt cgt     1968
Val Ala Pro Ala Ser Val Leu Asn Asn Trp Ala Asp Glu Ile Ser Arg
                645                 650                 655

ttc tgt cct gac ttg aaa act ctt cca tat tgg gga gga tta caa gaa     2016
Phe Cys Pro Asp Leu Lys Thr Leu Pro Tyr Trp Gly Gly Leu Gln Glu
            660                 665                 670

cga aca att tta aga aag aat atc aat ccc aag cgt atg tac cga agg     2064
Arg Thr Ile Leu Arg Lys Asn Ile Asn Pro Lys Arg Met Tyr Arg Arg
        675                 680                 685

gat gct ggc ttt cat att ttg att act agc tat cag cta tta gtc act     2112
Asp Ala Gly Phe His Ile Leu Ile Thr Ser Tyr Gln Leu Leu Val Thr
    690                 695                 700

gat gaa aag tat ttt cgc cgg gtg aag tgg caa tat atg gtg cta gat     2160
Asp Glu Lys Tyr Phe Arg Arg Val Lys Trp Gln Tyr Met Val Leu Asp
705                 710                 715                 720

gag gcc caa gca atc aag agt tcc tcc agt ata aga tgg aaa acc ctt     2208
Glu Ala Gln Ala Ile Lys Ser Ser Ser Ser Ile Arg Trp Lys Thr Leu
                725                 730                 735

ctt agt ttt aac tgt cgg aac cga ttg ctt ctg act ggt act cca att     2256
Leu Ser Phe Asn Cys Arg Asn Arg Leu Leu Leu Thr Gly Thr Pro Ile
            740                 745                 750

cag aac aac atg gca gag tta tgg gcc ctg ctg cat ttc atc atg cca     2304
Gln Asn Asn Met Ala Glu Leu Trp Ala Leu Leu His Phe Ile Met Pro
        755                 760                 765

atg ttg ttt gac aac cat gat caa ttt aat gaa tgg ttc tca aaa gga     2352
Met Leu Phe Asp Asn His Asp Gln Phe Asn Glu Trp Phe Ser Lys Gly
    770                 775                 780

att gag aat cat gct gaa cac gga ggc act tta aat gag cac cag ctt     2400
Ile Glu Asn His Ala Glu His Gly Gly Thr Leu Asn Glu His Gln Leu
785                 790                 795                 800

aac aga ctg cat gcg atc ttg aaa ccg ttc atg ctt cga cgg gta aaa     2448
Asn Arg Leu His Ala Ile Leu Lys Pro Phe Met Leu Arg Arg Val Lys
                805                 810                 815

aag gat gtg gtt tct gag cta act aca aag acg gaa gtt aca gta cac     2496
Lys Asp Val Val Ser Glu Leu Thr Thr Lys Thr Glu Val Thr Val His
            820                 825                 830

tgc aag ctc agt tct cga caa caa gct ttt tat cag gct att aag aac     2544
Cys Lys Leu Ser Ser Arg Gln Gln Ala Phe Tyr Gln Ala Ile Lys Asn
        835                 840                 845

aaa att tct ctg gct gag ttg ttt gat agc aac cgc gga caa ttt act     2592
Lys Ile Ser Leu Ala Glu Leu Phe Asp Ser Asn Arg Gly Gln Phe Thr
    850                 855                 860

gat aag aaa gta ttg aat tta atg aat att gtc att caa cta agg aag     2640
Asp Lys Lys Val Leu Asn Leu Met Asn Ile Val Ile Gln Leu Arg Lys
```

```
865                 870                 875                 880

gtt tgc aac cat cca gag ttg ttc gaa agg aat gaa ggg agc tcg tat     2688
Val Cys Asn His Pro Glu Leu Phe Glu Arg Asn Glu Gly Ser Ser Tyr
                885                 890                 895

ctc tac ttt gga gtg act tcc aat tct ctt ttg ccc cat ccc ttt ggt     2736
Leu Tyr Phe Gly Val Thr Ser Asn Ser Leu Leu Pro His Pro Phe Gly
            900                 905                 910

gag cta gag gat gta cat tat tct ggt ggt caa aat ccg ata ata tac     2784
Glu Leu Glu Asp Val His Tyr Ser Gly Gly Gln Asn Pro Ile Ile Tyr
        915                 920                 925

aag ata cct aag cta cta cac caa gag gtg ctc caa aat tct gaa aca     2832
Lys Ile Pro Lys Leu Leu His Gln Glu Val Leu Gln Asn Ser Glu Thr
    930                 935                 940

ttt tgt tct tct gtc ggg cgt ggc atc tca aga gaa tct ttt ctg aag     2880
Phe Cys Ser Ser Val Gly Arg Gly Ile Ser Arg Glu Ser Phe Leu Lys
945                 950                 955                 960

cat ttt aat ata tat tca cct gag tat att ctt aag tca ata ttc cca     2928
His Phe Asn Ile Tyr Ser Pro Glu Tyr Ile Leu Lys Ser Ile Phe Pro
                965                 970                 975

tct gat agt ggg gta gat caa gtg gtt agt gga agt gga gca ttt ggc     2976
Ser Asp Ser Gly Val Asp Gln Val Val Ser Gly Ser Gly Ala Phe Gly
            980                 985                 990

ttt tca cgc ttg atg gat cta tca  cca tca gaa gtt gga  tat ctg gct   3024
Phe Ser Arg Leu Met Asp Leu Ser  Pro Ser Glu Val Gly  Tyr Leu Ala
        995                 1000                1005

ctg tgt  tct gtt gca gaa agg  cta tta ttt tct ata  ctg agg tgg      3069
Leu Cys  Ser Val Ala Glu Arg  Leu Leu Phe Ser Ile  Leu Arg Trp
    1010                1015                1020

gag cgg  caa ttt ttg gat gaa  tta gtt aac tct ctt  atg gag tcc      3114
Glu Arg  Gln Phe Leu Asp Glu  Leu Val Asn Ser Leu  Met Glu Ser
    1025                1030                1035

aag gat  ggt gat ctt agt gac  aat aac atc gag aga  gtt aaa acc      3159
Lys Asp  Gly Asp Leu Ser Asp  Asn Asn Ile Glu Arg  Val Lys Thr
    1040                1045                1050

aaa gct  gtc aca aga atg ttg  ctg atg cca tca aaa  gtt gaa acg      3204
Lys Ala  Val Thr Arg Met Leu  Leu Met Pro Ser Lys  Val Glu Thr
    1055                1060                1065

aat ttt  cag aaa agg aga cta  agc aca ggg cct acc  cgt cct tca      3249
Asn Phe  Gln Lys Arg Arg Leu  Ser Thr Gly Pro Thr  Arg Pro Ser
    1070                1075                1080

ttt gaa  gcg cta gtg atc tct  cat cag gat agg ttt  ctt tca agt      3294
Phe Glu  Ala Leu Val Ile Ser  His Gln Asp Arg Phe  Leu Ser Ser
    1085                1090                1095

atc aaa  ctc ctg cat tct gca  tat act tat atc cca  aaa gcc aga      3339
Ile Lys  Leu Leu His Ser Ala  Tyr Thr Tyr Ile Pro  Lys Ala Arg
    1100                1105                1110

gct cca  cct gta agc att cat  tgc tcg gac aga aat  tcg gca tac      3384
Ala Pro  Pro Val Ser Ile His  Cys Ser Asp Arg Asn  Ser Ala Tyr
    1115                1120                1125
```

```
aga gtt  aca gaa gaa tta cat  caa cca tgg ctt aag  aga cta tta      3429
Arg Val  Thr Glu Glu Leu His  Gln Pro Trp Leu Lys  Arg Leu Leu
    1130                 1135                 1140

atc ggt  ttt gca cga acg tca  gaa gct aat gga ccc  agg aag cct      3474
Ile Gly  Phe Ala Arg Thr Ser  Glu Ala Asn Gly Pro  Arg Lys Pro
    1145                 1150                 1155

aac agc  ttt cca cat cct tta  atc caa gaa att gat  tca gaa ctt      3519
Asn Ser  Phe Pro His Pro Leu  Ile Gln Glu Ile Asp  Ser Glu Leu
    1160                 1165                 1170

cca gtt  gtg cag cct gcg ctt  caa ctg aca cac aga  ata ttt ggt      3564
Pro Val  Val Gln Pro Ala Leu  Gln Leu Thr His Arg  Ile Phe Gly
    1175                 1180                 1185

tct tgc  cct cca atg caa agt  ttt gac cca gca aag  ttg ctc acg      3609
Ser Cys  Pro Pro Met Gln Ser  Phe Asp Pro Ala Lys  Leu Leu Thr
    1190                 1195                 1200

gac tct  ggg aag ctg cag aca  ctt gat ata tta ttg  aag cgg ctt      3654
Asp Ser  Gly Lys Leu Gln Thr  Leu Asp Ile Leu Leu  Lys Arg Leu
    1205                 1210                 1215

cga gct  gga aat cac agg gtg  ctc ctg ttt gca caa  atg aca aag      3699
Arg Ala  Gly Asn His Arg Val  Leu Leu Phe Ala Gln  Met Thr Lys
    1220                 1225                 1230

atg ctg  aac att ctc gag gat  tat atg aac tat aga  aag tac aag      3744
Met Leu  Asn Ile Leu Glu Asp  Tyr Met Asn Tyr Arg  Lys Tyr Lys
    1235                 1240                 1245

tac ctc  agg ctt gat gga tcc  tcc acc atc atg gat  cgc cga gat      3789
Tyr Leu  Arg Leu Asp Gly Ser  Ser Thr Ile Met Asp  Arg Arg Asp
    1250                 1255                 1260

atg gtt  agg gat ttt cag cat  agg agc gat att ttt  gta ttc ttg      3834
Met Val  Arg Asp Phe Gln His  Arg Ser Asp Ile Phe  Val Phe Leu
    1265                 1270                 1275

ctg agc  acc aga gct gga gga  ctt ggt atc aac ttg  acg gct gca      3879
Leu Ser  Thr Arg Ala Gly Gly  Leu Gly Ile Asn Leu  Thr Ala Ala
    1280                 1285                 1290

gac act  gtc att ttc tat gaa  agt gat tgg aat ccc  acc ttg gat      3924
Asp Thr  Val Ile Phe Tyr Glu  Ser Asp Trp Asn Pro  Thr Leu Asp
    1295                 1300                 1305

tta caa  gct atg gac agg gct  cat cgt ctt gga cag  aca aaa gat      3969
Leu Gln  Ala Met Asp Arg Ala  His Arg Leu Gly Gln  Thr Lys Asp
    1310                 1315                 1320

gtt act  gtt tat cgt ctc atc  tgt aag gag acg gtg  gaa gag aaa      4014
Val Thr  Val Tyr Arg Leu Ile  Cys Lys Glu Thr Val  Glu Glu Lys
    1325                 1330                 1335

att ttg  cac agg gca agt cag  aaa aat aca gtt caa  cag ctt gtt      4059
Ile Leu  His Arg Ala Ser Gln  Lys Asn Thr Val Gln  Gln Leu Val
    1340                 1345                 1350

atg act  gga ggg cat gtt cag  ggt gat gat ttt ctt  gga gct gcg      4104
Met Thr  Gly Gly His Val Gln  Gly Asp Asp Phe Leu  Gly Ala Ala
    1355                 1360                 1365

gat gtg  gta tct ctg cta atg  gat gat gcg gag gca  gca caa ctg      4149
Asp Val  Val Ser Leu Leu Met  Asp Asp Ala Glu Ala  Ala Gln Leu
```

```
    1370                1375                1380

gag cag  aaa ttc aga gaa cta  cca tta cag gta aag  gac agg cag      4194
Glu Gln  Lys Phe Arg Glu Leu  Pro Leu Gln Val Lys  Asp Arg Gln
    1385                1390                1395

aag aaa  aag acg aaa cgt atc  aga ata gat gct gaa  gga gat gca      4239
Lys Lys  Lys Thr Lys Arg Ile  Arg Ile Asp Ala Glu  Gly Asp Ala
    1400                1405                1410

act ttg  gaa gag tta gaa gat  gtt gac cga cag gat  aac gga cag      4284
Thr Leu  Glu Glu Leu Glu Asp  Val Asp Arg Gln Asp  Asn Gly Gln
    1415                1420                1425

gaa cct  ttg gaa gaa ccg gaa  aag cca aaa tcc agt  aat aaa aag      4329
Glu Pro  Leu Glu Glu Pro Glu  Lys Pro Lys Ser Ser  Asn Lys Lys
    1430                1435                1440

agg aga  gct gct tca aat ccg  aaa gct aga gct cct  cag aaa gca      4374
Arg Arg  Ala Ala Ser Asn Pro  Lys Ala Arg Ala Pro  Gln Lys Ala
    1445                1450                1455

aag gaa  gaa gca aat ggt gaa  gat act cct cag agg  aca aaa agg      4419
Lys Glu  Glu Ala Asn Gly Glu  Asp Thr Pro Gln Arg  Thr Lys Arg
    1460                1465                1470

gta aag  aga caa aca aag agc  ata aac gaa agt ctt  gaa cct gta      4464
Val Lys  Arg Gln Thr Lys Ser  Ile Asn Glu Ser Leu  Glu Pro Val
    1475                1480                1485

ttc tct  gcc tct gta aca gaa  tca aat aaa gga ttc  gat cca agt      4509
Phe Ser  Ala Ser Val Thr Glu  Ser Asn Lys Gly Phe  Asp Pro Ser
    1490                1495                1500

agc tcc  gct aac taa                                            4524
Ser Ser  Ala Asn
    1505
```

<210> 2

<211> 1507

<212> PRT

<213> Arabidopsis thaliana

<400> 2

```
Met Asp Pro Ser Arg Arg Pro Pro Lys Asp Ser Pro Tyr Ala Asn Leu
1               5                  10                  15

Phe Asp Leu Glu Pro Leu Met Lys Phe Arg Ile Pro Lys Pro Glu Asp
            20                  25                  30

Glu Val Asp Tyr Tyr Gly Ser Ser Ser Gln Asp Glu Ser Arg Ser Thr
        35                  40                  45

Gln Gly Gly Val Val Ala Asn Tyr Ser Asn Gly Ser Lys Ser Arg Met
    50                  55                  60
```

Asn Ala Ser Ser Lys Lys Arg Lys Arg Trp Thr Glu Ala Glu Asp Ala
65 70 75 80

Glu Asp Asp Asp Asp Leu Tyr Asn Gln His Val Thr Glu Glu His Tyr
85 90 95

Arg Ser Met Leu Gly Glu His Val Gln Lys Phe Lys Asn Arg Ser Lys
100 105 110

Glu Thr Gln Gly Asn Pro Pro His Leu Met Gly Phe Pro Val Leu Lys
115 120 125

Ser Asn Val Gly Ser Tyr Arg Gly Arg Lys Pro Gly Asn Asp Tyr His
130 135 140

Gly Arg Phe Tyr Asp Met Asp Asn Ser Pro Asn Phe Ala Ala Asp Val
145 150 155 160

Thr Pro His Arg Arg Gly Ser Tyr His Asp Arg Asp Ile Thr Pro Lys
165 170 175

Ile Ala Tyr Glu Pro Ser Tyr Leu Asp Ile Gly Asp Gly Val Ile Tyr
180 185 190

Lys Ile Pro Pro Ser Tyr Asp Lys Leu Val Ala Ser Leu Asn Leu Pro
195 200 205

Ser Phe Ser Asp Ile His Val Glu Glu Phe Tyr Leu Lys Gly Thr Leu
210 215 220

Asp Leu Arg Ser Leu Ala Glu Leu Met Ala Ser Asp Lys Arg Ser Gly
225 230 235 240

Val Arg Ser Arg Asn Gly Met Gly Glu Pro Arg Pro Gln Tyr Glu Ser
245 250 255

Leu Gln Ala Arg Met Lys Ala Leu Ser Pro Ser Asn Ser Thr Pro Asn
260 265 270

Phe Ser Leu Lys Val Ser Glu Ala Ala Met Asn Ser Ala Ile Pro Glu
275 280 285

Gly Ser Ala Gly Ser Thr Ala Arg Thr Ile Leu Ser Glu Gly Gly Val
290 295 300

Leu Gln Val His Tyr Val Lys Ile Leu Glu Lys Gly Asp Thr Tyr Glu
305 310 315 320

Ile Val Lys Arg Ser Leu Pro Lys Lys Leu Lys Ala Lys Asn Asp Pro

325 330 335

Ala Val Ile Glu Lys Thr Glu Arg Asp Lys Ile Arg Lys Ala Trp Ile
340 345 350

Asn Ile Val Arg Arg Asp Ile Ala Lys His His Arg Ile Phe Thr Thr
355 360 365

Phe His Arg Lys Leu Ser Ile Asp Ala Lys Arg Phe Ala Asp Gly Cys
370 375 380

Gln Arg Glu Val Arg Met Lys Val Gly Arg Ser Tyr Lys Ile Pro Arg
385 390 395 400

Thr Ala Pro Ile Arg Thr Arg Lys Ile Ser Arg Asp Met Leu Leu Phe
405 410 415

Trp Lys Arg Tyr Asp Lys Gln Met Ala Glu Glu Arg Lys Lys Gln Glu
420 425 430

Lys Glu Ala Ala Glu Ala Phe Lys Arg Glu Gln Glu Gln Arg Glu Ser
435 440 445

Lys Arg Gln Gln Gln Arg Leu Asn Phe Leu Ile Lys Gln Thr Glu Leu
450 455 460

Tyr Ser His Phe Met Gln Asn Lys Thr Asp Ser Asn Pro Ser Glu Ala
465 470 475 480

Leu Pro Ile Gly Asp Glu Asn Pro Ile Asp Glu Val Leu Pro Glu Thr
485 490 495

Ser Ala Ala Glu Pro Ser Glu Val Glu Asp Pro Glu Glu Ala Glu Leu
500 505 510

Lys Glu Lys Val Leu Arg Ala Ala Gln Asp Ala Val Ser Lys Gln Lys
515 520 525

Gln Ile Thr Asp Ala Phe Asp Thr Glu Tyr Met Lys Leu Arg Gln Thr
530 535 540

Ser Glu Met Glu Gly Pro Leu Asn Asp Ile Ser Val Ser Gly Ser Ser
545 550 555 560

Asn Ile Asp Leu His Asn Pro Ser Thr Met Pro Val Thr Ser Thr Val
565 570 575

Gln Thr Pro Glu Leu Phe Lys Gly Thr Leu Lys Glu Tyr Gln Met Lys
580 585 590

Gly Leu Gln Trp Leu Val Asn Cys Tyr Glu Gln Gly Leu Asn Gly Ile
595 600 605

Leu Ala Asp Glu Met Gly Leu Gly Lys Thr Ile Gln Ala Met Ala Phe
610 615 620

Leu Ala His Leu Ala Glu Glu Lys Asn Ile Trp Gly Pro Phe Leu Val
625 630 635 640

Val Ala Pro Ala Ser Val Leu Asn Asn Trp Ala Asp Glu Ile Ser Arg
645 650 655

Phe Cys Pro Asp Leu Lys Thr Leu Pro Tyr Trp Gly Gly Leu Gln Glu
660 665 670

Arg Thr Ile Leu Arg Lys Asn Ile Asn Pro Lys Arg Met Tyr Arg Arg
675 680 685

Asp Ala Gly Phe His Ile Leu Ile Thr Ser Tyr Gln Leu Leu Val Thr
690 695 700

Asp Glu Lys Tyr Phe Arg Arg Val Lys Trp Gln Tyr Met Val Leu Asp
705 710 715 720

Glu Ala Gln Ala Ile Lys Ser Ser Ser Ser Ile Arg Trp Lys Thr Leu
725 730 735

Leu Ser Phe Asn Cys Arg Asn Arg Leu Leu Leu Thr Gly Thr Pro Ile
740 745 750

Gln Asn Asn Met Ala Glu Leu Trp Ala Leu Leu His Phe Ile Met Pro
755 760 765

Met Leu Phe Asp Asn His Asp Gln Phe Asn Glu Trp Phe Ser Lys Gly
770 775 780

Ile Glu Asn His Ala Glu His Gly Gly Thr Leu Asn Glu His Gln Leu
785 790 795 800

Asn Arg Leu His Ala Ile Leu Lys Pro Phe Met Leu Arg Arg Val Lys
805 810 815

Lys Asp Val Val Ser Glu Leu Thr Thr Lys Thr Glu Val Thr Val His
820 825 830

Cys Lys Leu Ser Ser Arg Gln Gln Ala Phe Tyr Gln Ala Ile Lys Asn
835 840 845

Lys Ile Ser Leu Ala Glu Leu Phe Asp Ser Asn Arg Gly Gln Phe Thr

```
    850                 855                 860

Asp Lys Lys Val Leu Asn Leu Met Asn Ile Val Ile Gln Leu Arg Lys
865                 870                 875                 880

Val Cys Asn His Pro Glu Leu Phe Glu Arg Asn Glu Gly Ser Ser Tyr
                885                 890                 895

Leu Tyr Phe Gly Val Thr Ser Asn Ser Leu Leu Pro His Pro Phe Gly
            900                 905                 910

Glu Leu Glu Asp Val His Tyr Ser Gly Gly Gln Asn Pro Ile Ile Tyr
        915                 920                 925

Lys Ile Pro Lys Leu Leu His Gln Glu Val Leu Gln Asn Ser Glu Thr
    930                 935                 940

Phe Cys Ser Ser Val Gly Arg Gly Ile Ser Arg Glu Ser Phe Leu Lys
945                 950                 955                 960

His Phe Asn Ile Tyr Ser Pro Glu Tyr Ile Leu Lys Ser Ile Phe Pro
                965                 970                 975

Ser Asp Ser Gly Val Asp Gln Val Val Ser Gly Ser Gly Ala Phe Gly
            980                 985                 990

Phe Ser Arg Leu Met Asp Leu Ser Pro Ser Glu Val Gly Tyr Leu Ala
        995                 1000                1005

Leu Cys Ser Val Ala Glu Arg Leu Leu Phe Ser Ile Leu Arg Trp
    1010                1015                1020

Glu Arg Gln Phe Leu Asp Glu Leu Val Asn Ser Leu Met Glu Ser
    1025                1030                1035

Lys Asp Gly Asp Leu Ser Asp Asn Asn Ile Glu Arg Val Lys Thr
    1040                1045                1050

Lys Ala Val Thr Arg Met Leu Leu Met Pro Ser Lys Val Glu Thr
    1055                1060                1065

Asn Phe Gln Lys Arg Arg Leu Ser Thr Gly Pro Thr Arg Pro Ser
    1070                1075                1080

Phe Glu Ala Leu Val Ile Ser His Gln Asp Arg Phe Leu Ser Ser
    1085                1090                1095

Ile Lys Leu Leu His Ser Ala Tyr Thr Tyr Ile Pro Lys Ala Arg
    1100                1105                1110
```

Ala Pro Pro Val Ser Ile His Cys Ser Asp Arg Asn Ser Ala Tyr
1115 1120 1125

Arg Val Thr Glu Glu Leu His Gln Pro Trp Leu Lys Arg Leu Leu
1130 1135 1140

Ile Gly Phe Ala Arg Thr Ser Glu Ala Asn Gly Pro Arg Lys Pro
1145 1150 1155

Asn Ser Phe Pro His Pro Leu Ile Gln Glu Ile Asp Ser Glu Leu
1160 1165 1170

Pro Val Val Gln Pro Ala Leu Gln Leu Thr His Arg Ile Phe Gly
1175 1180 1185

Ser Cys Pro Pro Met Gln Ser Phe Asp Pro Ala Lys Leu Leu Thr
1190 1195 1200

Asp Ser Gly Lys Leu Gln Thr Leu Asp Ile Leu Leu Lys Arg Leu
1205 1210 1215

Arg Ala Gly Asn His Arg Val Leu Leu Phe Ala Gln Met Thr Lys
1220 1225 1230

Met Leu Asn Ile Leu Glu Asp Tyr Met Asn Tyr Arg Lys Tyr Lys
1235 1240 1245

Tyr Leu Arg Leu Asp Gly Ser Ser Thr Ile Met Asp Arg Arg Asp
1250 1255 1260

Met Val Arg Asp Phe Gln His Arg Ser Asp Ile Phe Val Phe Leu
1265 1270 1275

Leu Ser Thr Arg Ala Gly Gly Leu Gly Ile Asn Leu Thr Ala Ala
1280 1285 1290

Asp Thr Val Ile Phe Tyr Glu Ser Asp Trp Asn Pro Thr Leu Asp
1295 1300 1305

Leu Gln Ala Met Asp Arg Ala His Arg Leu Gly Gln Thr Lys Asp
1310 1315 1320

Val Thr Val Tyr Arg Leu Ile Cys Lys Glu Thr Val Glu Glu Lys
1325 1330 1335

Ile Leu His Arg Ala Ser Gln Lys Asn Thr Val Gln Gln Leu Val
1340 1345 1350

Met Thr Gly Gly His Val Gln Gly Asp Asp Phe Leu Gly Ala Ala

1355 1360 1365

Asp Val Val Ser Leu Leu Met Asp Asp Ala Glu Ala Ala Gln Leu
1370 1375 1380

Glu Gln Lys Phe Arg Glu Leu Pro Leu Gln Val Lys Asp Arg Gln
1385 1390 1395

Lys Lys Lys Thr Lys Arg Ile Arg Ile Asp Ala Glu Gly Asp Ala
1400 1405 1410

Thr Leu Glu Glu Leu Glu Asp Val Asp Arg Gln Asp Asn Gly Gln
1415 1420 1425

Glu Pro Leu Glu Glu Pro Glu Lys Pro Lys Ser Ser Asn Lys Lys
1430 1435 1440

Arg Arg Ala Ala Ser Asn Pro Lys Ala Arg Ala Pro Gln Lys Ala
1445 1450 1455

Lys Glu Glu Ala Asn Gly Glu Asp Thr Pro Gln Arg Thr Lys Arg
1460 1465 1470

Val Lys Arg Gln Thr Lys Ser Ile Asn Glu Ser Leu Glu Pro Val
1475 1480 1485

Phe Ser Ala Ser Val Thr Glu Ser Asn Lys Gly Phe Asp Pro Ser
1490 1495 1500

Ser Ser Ala Asn
1505

<210> 3

<211> 4492

<212> DNA

<213> Arabidopsis thaliana

<400> 3
atggatcctt caagacgacc accgaaggac tctccttacg cgaatctatt cgatctcgag 60

ccgttgatga agtttagaat tccgaaacct gaagatgaag ttgattatta tgggagtagt 120

agccaggatg aaagtagaag cactcaaggt ggggtagtgg caaactacag caatgggtct 180

aaatcgagaa tgaatgcgag ctccaagaag agaaagcggt ggacagaagc tgaggatgca 240

gaggacgatg atgatctcta caatcaacat gttactgagg agcactaccg atcaatgctt 300

ggggagcatg tacaaaaatt caaaaatagg tccaaggaga ctcaagggaa tcctcctcat 360

ctgatgggtt ttccggtgct aaagagcaat gtgggcagtt acagaggtag gaaaccaggg 420
aatgattacc atgggaggtt ctatgacatg gacaactctc caaattttgc agctgatgtg 480
accccacata ggcgaggaag ctaccatgat cgtgatatta cacccaagat agcatatgaa 540
ccttcgtatt tggacattgg tgatggtgtc atctacaaaa tccccccaag ttatgacaag 600
ctggtggcat cattaaactt accgagcttt tcagacattc atgtggaaga attttacttg 660
aaaggaactc tggatctgag atcattagca gaactgatgg caagtgataa aaggtctgga 720
gtaagaagcc gtaatggaat gggtgagcct cgacctcaat atgaatctct tcaagctaga 780
atgaaggccc tgtcaccttc aaactccacc ccaaatttta gcctcaaggt gtcagaagct 840
gcaatgaatt ctgccattcc agaaggatct gctggaagta ctgcacggac aattctgtct 900
gagggtggtg ttttacaggt ccattacgtg aagattctgg agaaggggga tacatacgag 960
attgttaaac gaagtctacc gaagaagctg aaagcaaaga atgatcctgc agtcattgag 1020
aaaacagaaa gggataaaat tagaaaagcc tggatcaata ttgtcagaag agatatagca 1080
aaacaccata gaattttcac tacttttcat cgtaaactat caattgatgc caagaggttt 1140
gcagatggtt gccaaagaga ggtgagaatg aaggtgggta gatcatacaa aatcccaaga 1200
actgcaccaa ttcgcactag gaagatatcc agagacatgc tgctattctg gaagcgatat 1260
gacaagcaga tggcagaaga gaggaaaaag caagaaaagg aagctgcaga ggcttttaaa 1320
cgtgaacagg agcagcgaga gtcaaaaagg cagcaacaaa ggctcaattt ccttattaaa 1380
cagactgagc tttacagtca cttcatgcaa aacaagaccg attcgaatcc ttccgaagcc 1440
ttaccaatag gtgatgaaaa tccgattgac gaagtgctcc cagaaacttc agcggcagaa 1500
ccttctgagg tagaggatcc tgaagaggct gaactgaagg aaaaggtctt gagagctgcc 1560
caagatgcgg tgtctaagca gaagcaaata acagatgcat ttgacactga atatatgaag 1620
ctacgccaaa cttctgaaat ggaaggtcct ttaaatgata tatcagtttc tggctcgagc 1680
aatatagatt tgcataaccc atctacaatg cctgttacat caacagttca gactccagag 1740
ttatttaaag gaacccttaa agaataccaa atgaaaggcc ttcagtggct agtcaattgt 1800
tatgagcagg gtttgaatgg catacttgct gatgaaatgg gcttgggtaa gactattcaa 1860
gctatggcgt tcttggcaca tttggctgag gaaaagaaca tttggggtcc atttcttgtt 1920
gttgcccctg cctctgttct taacaattgg gctgatgaaa tcagtcgttt ctgtcctgac 1980
ttgaaaactc ttccatattg gggaggatta caagaacgaa caattttaag aaagaatatc 2040
aatcccaagc gtatgtaccg aagggatgct ggctttcata ttttgattac tagctatcag 2100
ctattagtca ctgatgaaaa gtattttcgc cgggtgaagt ggcaatatat ggtgctagat 2160
gaggcccaag caatcaagag ttcctccagt ataagatgga aaacccttct tagttttaac 2220
tgtcggaacc gattgcttct gactggtact ccaattcaga acaacatggc agagttatgg 2280
gccctgctgc atttcatcat gccaatgttg tttgacaacc atgatcaatt taatgaatgg 2340

ttctcaaaag gaattgagaa tcatgctgaa cacggaggca ctttaaatga gcaccagctt 2400
aacagactgc atgcgatctt gaaaccgttc atgcttcgac gggtaaaaaa ggatgtggtt 2460
tctgagctaa ctacaaagac ggaagttaca gtacactgca agctcagttc tcgacaacaa 2520
gctttttatc aggctattaa gaacaaaatt tctctggctg agttgtttga tagcaaccgc 2580
ggacaattta ctgataagaa agtattgaat ttaatgaata ttgtcattca actaaggaag 2640
gtttgcaacc atccagagtt gttcgaaagg aatgaaggga gctcgtatct ctactttgga 2700
gtgacttcca attctctttt gccccatccc tttggtgagc tagaggatgt acattattct 2760
ggtggtcaaa atccgataat atacaagata cctaagctac tacaccaaga ggtgctccaa 2820
aattctgaaa cattttgttc ttctgtcggg cgtggcatct caagagaatc ttttctgaag 2880
cattttaata tatattcacc tgagtatatt cttaagtcaa tattcccatc tgatagtggg 2940
gtagatcaag tggttagtgg aagtggagca tttggctttt cacgcttgat ggatctatca 3000
ccatcagaag ttggatatct ggctctgtgt tctgttgcag aaaggctatt attttctata 3060
ctgaggtggg agcggcaatt tttggatgaa ttagttaact ctcttatgga gtccaaggat 3120
ggtgatctta gtgacaataa catcgagaga gttaaaacca aagctgtcac aagaatgttg 3180
ctgatgccat caaaagttga aacgaatttt cagaaaagga gactaagcac agggcctacc 3240
cgtccttcat ttgaagcgct agtgatctct catcaggata ggtttctttc aagtatcaaa 3300
ctcctgcatt ctgcatatac ttatatccca aaagccagag ctccacctgt aagcattcat 3360
tgctcggaca gaaattcggc atacagagtt acagaagaat tacatcaacc atggcttaag 3420
agactattaa tcggttttgc acgaacgtca gaagctaatg gacccaggaa gcctaacagc 3480
tttccacatc ctttaatcca agaaattgat tcagaacttc cagttgtgca gcctgcgctt 3540
caactgacac acagaatatt tggttcttgc cctccaatgc aaagttttga cccagcaaag 3600
ttgctcacgg actctgggaa gctgcagaca cttgatatat tattgaagcg gcttcgagct 3660
ggaaatcaca gggtgctcct gtttgcacaa atgacaaaga tgctgaacat tctcgaggat 3720
tatatgaact atagaaagta caagtacctc aggcttgatg gatcctccac catcatggat 3780
cgccgagata tggttaggga ttttcagcat aggagcgata tttttgtatt cttgctgagc 3840
accagagctg gaggacttgg tatcaacttg acggctgcag acactgtcat tttctatgaa 3900
agtgattgga atcccacctt ggatttacaa gctatggaca gggctcatcg tcttggacag 3960
acaaaagatg gagacggtgg aagagaaaat tttgcacagg gcaagtcaga aaaatacagt 4020
tcaacagctt gttatgactg gagggcatgt tcagggtgat gattttcttg gagctgcgga 4080
tgtggtatct ctgctaatgg atgatgcgga ggcagcacaa ctggagcaga aattcagaga 4140
actaccatta caggacaggc agaagaaaaa gacgaaacgt atcagaatag atgctgaagg 4200
agatgcaact ttggaagagt tagaagatgt tgaccgacag gataacggac aggaaccttt 4260
ggaagaaccg gaaaagccaa aatccagtaa taaaaagagg agagctgctt caaatccgaa 4320

agctagagct cctcagaaag caaaggaaga agcaaatggt gaagatactc ctcagaggac 4380

aaaaagggta aagagacaaa caaagagcat aaacgaaagt cttgaacctg tattctctgc 4440

ctctgtaaca gaatcaaata aaggattcga tccaagtagc tccgctaact aa 4492

<210> 4

<211> 1377

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(1377)

<223>

<400> 4

```
atg gag aaa gta aag att gaa gaa att cag tcc acc gct aag aaa caa      48
Met Glu Lys Val Lys Ile Glu Glu Ile Gln Ser Thr Ala Lys Lys Gln
1               5                  10                  15

cgg att gct act cac acc cat atc aaa ggc ctt ggc ctc gag cca act      96
Arg Ile Ala Thr His Thr His Ile Lys Gly Leu Gly Leu Glu Pro Thr
            20                  25                  30

ggt atc cct ata aaa ttg gca gct gga ttt gtt ggt caa ctt gag gct     144
Gly Ile Pro Ile Lys Leu Ala Ala Gly Phe Val Gly Gln Leu Glu Ala
        35                  40                  45

aga gag gca gct ggt ctt gta gtt gac atg att aag cag aag aaa atg     192
Arg Glu Ala Ala Gly Leu Val Val Asp Met Ile Lys Gln Lys Lys Met
    50                  55                  60

gcg ggc aag gct ctt ttg ctt gct gga cct cct gga act ggg aaa aca     240
Ala Gly Lys Ala Leu Leu Leu Ala Gly Pro Pro Gly Thr Gly Lys Thr
65                  70                  75                  80

gct ttg gct ctt gga atc tct caa gag ctg gga agc aag gtt cca ttc     288
Ala Leu Ala Leu Gly Ile Ser Gln Glu Leu Gly Ser Lys Val Pro Phe
                85                  90                  95

tgt cca atg gtt gga tct gag gtt tac tca tca gag gtt aag aaa aca     336
Cys Pro Met Val Gly Ser Glu Val Tyr Ser Ser Glu Val Lys Lys Thr
            100                 105                 110

gag gtt ctc atg gag aat ttt aga cgt gcc att ggt cta cgt atc aag     384
Glu Val Leu Met Glu Asn Phe Arg Arg Ala Ile Gly Leu Arg Ile Lys
        115                 120                 125

gaa acc aaa gaa gtc tat gaa ggg gag gtc acc gag ctg tca cca gaa     432
Glu Thr Lys Glu Val Tyr Glu Gly Glu Val Thr Glu Leu Ser Pro Glu
    130                 135                 140

gaa act gaa agc ctc act gga ggt tat ggt aaa agc atc agc cat gtt     480
Glu Thr Glu Ser Leu Thr Gly Gly Tyr Gly Lys Ser Ile Ser His Val
```

```
145                 150                 155                 160

gta att aca ctc aag aca gtc aaa gga acc aaa cat ctg aaa ttg gat      528
Val Ile Thr Leu Lys Thr Val Lys Gly Thr Lys His Leu Lys Leu Asp
                165                 170                 175

ccc act atc tat gat gcc ttg att aag gaa aag gta gct gta gga gat      576
Pro Thr Ile Tyr Asp Ala Leu Ile Lys Glu Lys Val Ala Val Gly Asp
            180                 185                 190

gta atc tat atc gaa gca aac agt gga gct gtc aaa cgg gta ggt aga      624
Val Ile Tyr Ile Glu Ala Asn Ser Gly Ala Val Lys Arg Val Gly Arg
        195                 200                 205

agt gat gct ttt gcc act gaa ttt gat ctg gaa gca gaa gaa tat gtt      672
Ser Asp Ala Phe Ala Thr Glu Phe Asp Leu Glu Ala Glu Glu Tyr Val
    210                 215                 220

cca ctt ccc aaa gga gag gtc cac aaa aag aaa gag ata gtg cag gat      720
Pro Leu Pro Lys Gly Glu Val His Lys Lys Lys Glu Ile Val Gln Asp
225                 230                 235                 240

gtc aca ctc caa gat ctg gat gca gca aat gct cga cct caa ggt ggc      768
Val Thr Leu Gln Asp Leu Asp Ala Ala Asn Ala Arg Pro Gln Gly Gly
                245                 250                 255

cag gat ata ctt tct ttg atg ggc caa atg atg aaa ccg cgg aag act      816
Gln Asp Ile Leu Ser Leu Met Gly Gln Met Met Lys Pro Arg Lys Thr
            260                 265                 270

gag atc act gat aag ctt cgg caa gaa att aac aag gtt gtg aac cga      864
Glu Ile Thr Asp Lys Leu Arg Gln Glu Ile Asn Lys Val Val Asn Arg
        275                 280                 285

tat ata gat gaa ggt gtg gca gag ctt gtt cca gga gtt cta ttt att      912
Tyr Ile Asp Glu Gly Val Ala Glu Leu Val Pro Gly Val Leu Phe Ile
    290                 295                 300

gat gag gtt cat atg ctt gat atg gag tgc ttc tca tac ttg aac cgt      960
Asp Glu Val His Met Leu Asp Met Glu Cys Phe Ser Tyr Leu Asn Arg
305                 310                 315                 320

gct ctt gag agc tca tta tct ccg ata gtg ata ttt gca aca aat aga     1008
Ala Leu Glu Ser Ser Leu Ser Pro Ile Val Ile Phe Ala Thr Asn Arg
                325                 330                 335

ggt gtt tgc aac gta aga ggg act gat atg ccc agc ccc cat gga gtc     1056
Gly Val Cys Asn Val Arg Gly Thr Asp Met Pro Ser Pro His Gly Val
            340                 345                 350

cct att gat cta tta gat cga ttg gtt atc atc cgg act caa atc tat     1104
Pro Ile Asp Leu Leu Asp Arg Leu Val Ile Ile Arg Thr Gln Ile Tyr
        355                 360                 365

gat ccc tct gaa atg atc cag att ata gcc att cgt gcg caa gtt gaa     1152
Asp Pro Ser Glu Met Ile Gln Ile Ile Ala Ile Arg Ala Gln Val Glu
    370                 375                 380

gaa tta acc gtg gat gaa gaa tgc ttg gtt cta ctt ggg gag att ggg     1200
Glu Leu Thr Val Asp Glu Glu Cys Leu Val Leu Leu Gly Glu Ile Gly
385                 390                 395                 400

caa aga act tca cta agg cac gct gtg cag ctt ctg tct cct gcc agc     1248
Gln Arg Thr Ser Leu Arg His Ala Val Gln Leu Leu Ser Pro Ala Ser
                405                 410                 415
```

```
att gta gcg aaa atg aat ggc cgt gac aat att tgc aag gct gat ata     1296
Ile Val Ala Lys Met Asn Gly Arg Asp Asn Ile Cys Lys Ala Asp Ile
            420                 425                 430

gag gaa gta aca tca ctc tac ttg gat gct aaa tct tca gca aag ctt     1344
Glu Glu Val Thr Ser Leu Tyr Leu Asp Ala Lys Ser Ser Ala Lys Leu
        435                 440                 445

ttg cat gag caa caa gaa aaa tac atc tca tga                        1377
Leu His Glu Gln Gln Glu Lys Tyr Ile Ser
    450                 455
```

<210> 5

<211> 458

<212> PRT

<213> Arabidopsis thaliana

<400> 5

```
Met Glu Lys Val Lys Ile Glu Glu Ile Gln Ser Thr Ala Lys Lys Gln
1               5                   10                  15

Arg Ile Ala Thr His Thr His Ile Lys Gly Leu Gly Leu Glu Pro Thr
            20                  25                  30

Gly Ile Pro Ile Lys Leu Ala Ala Gly Phe Val Gly Gln Leu Glu Ala
        35                  40                  45

Arg Glu Ala Ala Gly Leu Val Val Asp Met Ile Lys Gln Lys Lys Met
    50                  55                  60

Ala Gly Lys Ala Leu Leu Leu Ala Gly Pro Pro Gly Thr Gly Lys Thr
65                  70                  75                  80

Ala Leu Ala Leu Gly Ile Ser Gln Glu Leu Gly Ser Lys Val Pro Phe
                85                  90                  95

Cys Pro Met Val Gly Ser Glu Val Tyr Ser Ser Glu Val Lys Lys Thr
            100                 105                 110

Glu Val Leu Met Glu Asn Phe Arg Arg Ala Ile Gly Leu Arg Ile Lys
        115                 120                 125

Glu Thr Lys Glu Val Tyr Glu Gly Glu Val Thr Glu Leu Ser Pro Glu
    130                 135                 140

Glu Thr Glu Ser Leu Thr Gly Gly Tyr Gly Lys Ser Ile Ser His Val
145                 150                 155                 160

Val Ile Thr Leu Lys Thr Val Lys Gly Thr Lys His Leu Lys Leu Asp
```

```
                165                 170                 175

Pro Thr Ile Tyr Asp Ala Leu Ile Lys Glu Lys Val Ala Val Gly Asp
            180                 185                 190

Val Ile Tyr Ile Glu Ala Asn Ser Gly Ala Val Lys Arg Val Gly Arg
        195                 200                 205

Ser Asp Ala Phe Ala Thr Glu Phe Asp Leu Glu Ala Glu Glu Tyr Val
    210                 215                 220

Pro Leu Pro Lys Gly Glu Val His Lys Lys Lys Glu Ile Val Gln Asp
225                 230                 235                 240

Val Thr Leu Gln Asp Leu Asp Ala Ala Asn Ala Arg Pro Gln Gly Gly
                245                 250                 255

Gln Asp Ile Leu Ser Leu Met Gly Gln Met Met Lys Pro Arg Lys Thr
            260                 265                 270

Glu Ile Thr Asp Lys Leu Arg Gln Glu Ile Asn Lys Val Val Asn Arg
        275                 280                 285

Tyr Ile Asp Glu Gly Val Ala Glu Leu Val Pro Gly Val Leu Phe Ile
    290                 295                 300

Asp Glu Val His Met Leu Asp Met Glu Cys Phe Ser Tyr Leu Asn Arg
305                 310                 315                 320

Ala Leu Glu Ser Ser Leu Ser Pro Ile Val Ile Phe Ala Thr Asn Arg
                325                 330                 335

Gly Val Cys Asn Val Arg Gly Thr Asp Met Pro Ser Pro His Gly Val
            340                 345                 350

Pro Ile Asp Leu Leu Asp Arg Leu Val Ile Ile Arg Thr Gln Ile Tyr
        355                 360                 365

Asp Pro Ser Glu Met Ile Gln Ile Ile Ala Ile Arg Ala Gln Val Glu
    370                 375                 380

Glu Leu Thr Val Asp Glu Glu Cys Leu Val Leu Leu Gly Glu Ile Gly
385                 390                 395                 400

Gln Arg Thr Ser Leu Arg His Ala Val Gln Leu Leu Ser Pro Ala Ser
                405                 410                 415

Ile Val Ala Lys Met Asn Gly Arg Asp Asn Ile Cys Lys Ala Asp Ile
            420                 425                 430
```

Glu Glu Val Thr Ser Leu Tyr Leu Asp Ala Lys Ser Ser Ala Lys Leu
435 440 445

Leu His Glu Gln Gln Glu Lys Tyr Ile Ser
450 455

<210> 6

<211> 1584

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(1410)

<223>

<400> 6

```
atg gcg gaa cta aag cta tca gag agt cgg gac tta acc aga gtc gag      48
Met Ala Glu Leu Lys Leu Ser Glu Ser Arg Asp Leu Thr Arg Val Glu
1               5                  10                  15

cga atc ggc gca cac tca cac atc aga gga cta ggt ctc gac tct gcc      96
Arg Ile Gly Ala His Ser His Ile Arg Gly Leu Gly Leu Asp Ser Ala
            20                  25                  30

ctc gag ccg cga gct gtt tcc gaa ggt atg gtc ggt caa gtg aag gcg     144
Leu Glu Pro Arg Ala Val Ser Glu Gly Met Val Gly Gln Val Lys Ala
        35                  40                  45

cgt aaa gcc gcc ggt gta atc ctt cag atg att aga gaa ggg aaa atc     192
Arg Lys Ala Ala Gly Val Ile Leu Gln Met Ile Arg Glu Gly Lys Ile
    50                  55                  60

gcg ggt cgg gct att cta ata gcg ggt caa ccc gga acg ggt aag aca     240
Ala Gly Arg Ala Ile Leu Ile Ala Gly Gln Pro Gly Thr Gly Lys Thr
65                  70                  75                  80

gcg att gca atg ggt atg gcg aaa tct ctt ggc ttg gaa act cct ttt     288
Ala Ile Ala Met Gly Met Ala Lys Ser Leu Gly Leu Glu Thr Pro Phe
                85                  90                  95

gcg atg att gca gga agt gaa att ttc tca tta gag atg tca aag aca     336
Ala Met Ile Ala Gly Ser Glu Ile Phe Ser Leu Glu Met Ser Lys Thr
            100                 105                 110

gaa gct ttg act cag tct ttt cgt aaa gcg att ggt gtt agg atc aaa     384
Glu Ala Leu Thr Gln Ser Phe Arg Lys Ala Ile Gly Val Arg Ile Lys
        115                 120                 125

gaa gag aca gag gtt att gaa gga gaa gtt gtt gag gtt cag att gat     432
Glu Glu Thr Glu Val Ile Glu Gly Glu Val Val Glu Val Gln Ile Asp
    130                 135                 140
```

```
agg cct gct tct tct ggt gtt gct tcc aag tca ggg aag atg act atg      480
Arg Pro Ala Ser Ser Gly Val Ala Ser Lys Ser Gly Lys Met Thr Met
145                 150                 155                 160

aaa acg act gat atg gaa act gtg tat gat atg gga gct aag atg att      528
Lys Thr Thr Asp Met Glu Thr Val Tyr Asp Met Gly Ala Lys Met Ile
                165                 170                 175

gag gct ttg aac aag gag aaa gtg cag agt ggt gat gtt att gcc att      576
Glu Ala Leu Asn Lys Glu Lys Val Gln Ser Gly Asp Val Ile Ala Ile
            180                 185                 190

gat aaa gct act ggg aag att act aag ctt gga aga tcg ttt tcg agg      624
Asp Lys Ala Thr Gly Lys Ile Thr Lys Leu Gly Arg Ser Phe Ser Arg
        195                 200                 205

tct cgt gat tat gat gct atg ggt gcg cag acc aag ttt gtg cag tgc      672
Ser Arg Asp Tyr Asp Ala Met Gly Ala Gln Thr Lys Phe Val Gln Cys
    210                 215                 220

cct gaa ggt gag ttg cag aag agg aaa gag gtt gta cat tgt gtc act      720
Pro Glu Gly Glu Leu Gln Lys Arg Lys Glu Val Val His Cys Val Thr
225                 230                 235                 240

ctt cac gag att gat gtt atc aac agc agg aca caa ggg ttt ctt gcc      768
Leu His Glu Ile Asp Val Ile Asn Ser Arg Thr Gln Gly Phe Leu Ala
                245                 250                 255

ctt ttc act ggc gat act gga gaa atc cga tca gaa gtc cgg gaa caa      816
Leu Phe Thr Gly Asp Thr Gly Glu Ile Arg Ser Glu Val Arg Glu Gln
            260                 265                 270

att gat aca aaa gta gct gaa tgg aga gaa gaa gga aaa gca gag ata      864
Ile Asp Thr Lys Val Ala Glu Trp Arg Glu Glu Gly Lys Ala Glu Ile
        275                 280                 285

gtt ccc gga gtt ctc ttc att gat gaa gtc cac atg ctc gac atc gaa      912
Val Pro Gly Val Leu Phe Ile Asp Glu Val His Met Leu Asp Ile Glu
    290                 295                 300

tgc ttc tca ttc ctt aac cga gct cta gaa aac gaa atg tca cca atc      960
Cys Phe Ser Phe Leu Asn Arg Ala Leu Glu Asn Glu Met Ser Pro Ile
305                 310                 315                 320

ctt gtg gtg gca aca aac cga gga gtg acg aca atc cgt ggc aca aac     1008
Leu Val Val Ala Thr Asn Arg Gly Val Thr Thr Ile Arg Gly Thr Asn
                325                 330                 335

cag aaa tca cca cac ggg atc ccg att gat ctc ctt gac cgt ctt ctc     1056
Gln Lys Ser Pro His Gly Ile Pro Ile Asp Leu Leu Asp Arg Leu Leu
            340                 345                 350

atc atc act acc caa cct tac aca gac gat gac ata agg aag ata tta     1104
Ile Ile Thr Thr Gln Pro Tyr Thr Asp Asp Asp Ile Arg Lys Ile Leu
        355                 360                 365

gaa atc cgt tgc caa gag gaa gac gtt gag atg aac gaa gag gcc aaa     1152
Glu Ile Arg Cys Gln Glu Glu Asp Val Glu Met Asn Glu Glu Ala Lys
    370                 375                 380

cag ctt ttg aca ttg atc gga cgt gat aca tct cta agg tat gcg att     1200
Gln Leu Leu Thr Leu Ile Gly Arg Asp Thr Ser Leu Arg Tyr Ala Ile
385                 390                 395                 400

cat ctt ata acc gca gct gca ttg tca tgc cag aaa cgg aaa ggg aaa     1248
His Leu Ile Thr Ala Ala Ala Leu Ser Cys Gln Lys Arg Lys Gly Lys
```

```
                405                 410                 415

gtc gtg gag gtt gag gat att cag aga gtt tac aga ctg ttc ttg gat     1296
Val Val Glu Val Glu Asp Ile Gln Arg Val Tyr Arg Leu Phe Leu Asp
            420                 425                 430

gtg agg aga tcg atg cag tat ctt gtt gag tat cag agt cag tat atg     1344
Val Arg Arg Ser Met Gln Tyr Leu Val Glu Tyr Gln Ser Gln Tyr Met
        435                 440                 445

ttc agt gaa cca atc aaa aac gat gaa gct gct gca gaa gac gaa caa     1392
Phe Ser Glu Pro Ile Lys Asn Asp Glu Ala Ala Ala Glu Asp Glu Gln
    450                 455                 460

gat gct atg cag atc tga ggatccacct ctgtttgcct tatttatcat           1440
Asp Ala Met Gln Ile
465

gtttcgtggt gatatgtatg attaggatgt tgaactcgga tttatgtttt ttttttttta   1500

agttgtgacg agattcggtt ctagaaaatg atttaaccaa gttcaataca gatcggtttg   1560

gtacaaaaca aaaaaaaaaa aaaa                                          1584
```

<210> 7

<211> 469

<212> PRT

<213> Arabidopsis thaliana

<400> 7

```
Met Ala Glu Leu Lys Leu Ser Glu Ser Arg Asp Leu Thr Arg Val Glu
1               5                   10                  15

Arg Ile Gly Ala His Ser His Ile Arg Gly Leu Gly Leu Asp Ser Ala
            20                  25                  30

Leu Glu Pro Arg Ala Val Ser Glu Gly Met Val Gly Gln Val Lys Ala
        35                  40                  45

Arg Lys Ala Ala Gly Val Ile Leu Gln Met Ile Arg Glu Gly Lys Ile
    50                  55                  60

Ala Gly Arg Ala Ile Leu Ile Ala Gly Gln Pro Gly Thr Gly Lys Thr
65                  70                  75                  80

Ala Ile Ala Met Gly Met Ala Lys Ser Leu Gly Leu Glu Thr Pro Phe
                85                  90                  95

Ala Met Ile Ala Gly Ser Glu Ile Phe Ser Leu Glu Met Ser Lys Thr
            100                 105                 110

Glu Ala Leu Thr Gln Ser Phe Arg Lys Ala Ile Gly Val Arg Ile Lys
```

```
        115                 120                 125

Glu Glu Thr Glu Val Ile Glu Gly Glu Val Val Glu Val Gln Ile Asp
    130                 135                 140

Arg Pro Ala Ser Ser Gly Val Ala Ser Lys Ser Gly Lys Met Thr Met
145                 150                 155                 160

Lys Thr Thr Asp Met Glu Thr Val Tyr Asp Met Gly Ala Lys Met Ile
                165                 170                 175

Glu Ala Leu Asn Lys Glu Lys Val Gln Ser Gly Asp Val Ile Ala Ile
            180                 185                 190

Asp Lys Ala Thr Gly Lys Ile Thr Lys Leu Gly Arg Ser Phe Ser Arg
        195                 200                 205

Ser Arg Asp Tyr Asp Ala Met Gly Ala Gln Thr Lys Phe Val Gln Cys
    210                 215                 220

Pro Glu Gly Glu Leu Gln Lys Arg Lys Glu Val Val His Cys Val Thr
225                 230                 235                 240

Leu His Glu Ile Asp Val Ile Asn Ser Arg Thr Gln Gly Phe Leu Ala
                245                 250                 255

Leu Phe Thr Gly Asp Thr Gly Glu Ile Arg Ser Glu Val Arg Glu Gln
            260                 265                 270

Ile Asp Thr Lys Val Ala Glu Trp Arg Glu Glu Gly Lys Ala Glu Ile
        275                 280                 285

Val Pro Gly Val Leu Phe Ile Asp Glu Val His Met Leu Asp Ile Glu
    290                 295                 300

Cys Phe Ser Phe Leu Asn Arg Ala Leu Glu Asn Glu Met Ser Pro Ile
305                 310                 315                 320

Leu Val Val Ala Thr Asn Arg Gly Val Thr Thr Ile Arg Gly Thr Asn
                325                 330                 335

Gln Lys Ser Pro His Gly Ile Pro Ile Asp Leu Leu Asp Arg Leu Leu
            340                 345                 350

Ile Ile Thr Thr Gln Pro Tyr Thr Asp Asp Asp Ile Arg Lys Ile Leu
        355                 360                 365

Glu Ile Arg Cys Gln Glu Glu Asp Val Glu Met Asn Glu Glu Ala Lys
    370                 375                 380
```

Gln Leu Leu Thr Leu Ile Gly Arg Asp Thr Ser Leu Arg Tyr Ala Ile
385 390 395 400

His Leu Ile Thr Ala Ala Ala Leu Ser Cys Gln Lys Arg Lys Gly Lys
405 410 415

Val Val Glu Val Glu Asp Ile Gln Arg Val Tyr Arg Leu Phe Leu Asp
420 425 430

Val Arg Arg Ser Met Gln Tyr Leu Val Glu Tyr Gln Ser Gln Tyr Met
435 440 445

Phe Ser Glu Pro Ile Lys Asn Asp Glu Ala Ala Ala Glu Asp Glu Gln
450 455 460

Asp Ala Met Gln Ile
465

<210> 8

<211> 1422

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (64)..(1389)

<223>

<400> 8
atggcagaac taaggttatc agaaactcga gacttaacta ggatcgaaag aatcggagca 60

cac tca cac ata cga ggt tta ggt ctc gac tca gta ctc gag cca cga 108
Ser His Ile Arg Gly Leu Gly Leu Asp Ser Val Leu Glu Pro Arg
1 5 10 15

gcc gta tcc gaa gga atg gtt ggt caa atc aaa gca cgt aaa gcc gcc 156
Ala Val Ser Glu Gly Met Val Gly Gln Ile Lys Ala Arg Lys Ala Ala
20 25 30

gga gta acc ctc gag ttg atc aga gac ggc aaa atc tcg ggt cgg gct 204
Gly Val Thr Leu Glu Leu Ile Arg Asp Gly Lys Ile Ser Gly Arg Ala
35 40 45

ata ctt ata gcg ggt caa ccc gga acg ggt aaa atc gca ata gca atg 252
Ile Leu Ile Ala Gly Gln Pro Gly Thr Gly Lys Ile Ala Ile Ala Met
50 55 60

ggt ata gca aaa tca ctt gga caa gaa aca cca ttc act atg att gca 300
Gly Ile Ala Lys Ser Leu Gly Gln Glu Thr Pro Phe Thr Met Ile Ala

```
    65                  70                  75

gga agt gag atc ttt tct tta gag atg tca aag act gaa gct tta act      348
Gly Ser Glu Ile Phe Ser Leu Glu Met Ser Lys Thr Glu Ala Leu Thr
80                  85                  90                  95

caa gct ttt cgt aaa gct att ggt gtt agg atc aaa gaa gag act gac      396
Gln Ala Phe Arg Lys Ala Ile Gly Val Arg Ile Lys Glu Glu Thr Asp
                100                 105                 110

gtg ata gaa gga gaa gtt gtg acg att tcg att gat aga cct gct tct      444
Val Ile Glu Gly Glu Val Val Thr Ile Ser Ile Asp Arg Pro Ala Ser
            115                 120                 125

tct ggt ggt tct gtg aag aag act ggg aag ata aca atg aag acg act      492
Ser Gly Gly Ser Val Lys Lys Thr Gly Lys Ile Thr Met Lys Thr Thr
        130                 135                 140

gat atg gaa tct aat ttt gat ttg gga tgg aaa ttg att gag cca ttg      540
Asp Met Glu Ser Asn Phe Asp Leu Gly Trp Lys Leu Ile Glu Pro Leu
    145                 150                 155

gat aag gag aaa gta cag agt ggt gat gtt att gtt ttg gat agg ttt      588
Asp Lys Glu Lys Val Gln Ser Gly Asp Val Ile Val Leu Asp Arg Phe
160                 165                 170                 175

tgt ggg aag att act aag ctt gga aga tct ttt acg agg tct aga gat      636
Cys Gly Lys Ile Thr Lys Leu Gly Arg Ser Phe Thr Arg Ser Arg Asp
                180                 185                 190

ttt gat gtt atg ggt tca aag act aag ttt gtg cag tgc cct gaa ggt      684
Phe Asp Val Met Gly Ser Lys Thr Lys Phe Val Gln Cys Pro Glu Gly
            195                 200                 205

gag ctt gag aag agg aag gag gtt ttg cat tct gtc aca ctt cat gag      732
Glu Leu Glu Lys Arg Lys Glu Val Leu His Ser Val Thr Leu His Glu
        210                 215                 220

att gat gtt att aat agc agg act caa ggg tat cta gcc ctc ttc aca      780
Ile Asp Val Ile Asn Ser Arg Thr Gln Gly Tyr Leu Ala Leu Phe Thr
    225                 230                 235

ggt gat aca ggc gag att cgt tca gaa acc cga gag caa agc gat act      828
Gly Asp Thr Gly Glu Ile Arg Ser Glu Thr Arg Glu Gln Ser Asp Thr
240                 245                 250                 255

aaa gtg gca gag tgg aga gaa gaa ggg aaa gct gaa ata gta cct ggt      876
Lys Val Ala Glu Trp Arg Glu Glu Gly Lys Ala Glu Ile Val Pro Gly
                260                 265                 270

gtt ctc ttc att gat gaa gtc cat atg ctt gat atc gaa tgc ttc tct      924
Val Leu Phe Ile Asp Glu Val His Met Leu Asp Ile Glu Cys Phe Ser
            275                 280                 285

ttc ctg aat aga gct ctc gaa aac gat atg tca cca atc ctg gtc gtg      972
Phe Leu Asn Arg Ala Leu Glu Asn Asp Met Ser Pro Ile Leu Val Val
        290                 295                 300

gct aca aac aga gga atg aca aca atc cga gga aca aac cag ata tca     1020
Ala Thr Asn Arg Gly Met Thr Thr Ile Arg Gly Thr Asn Gln Ile Ser
    305                 310                 315

gca cat ggg atc cca atc gat ttt ctt gac cgt ctt ctt att atc aca     1068
Ala His Gly Ile Pro Ile Asp Phe Leu Asp Arg Leu Leu Ile Ile Thr
320                 325                 330                 335
```

```
aca cag cct tac aca caa gac gag atc aga aat att tta gag atc cgt     1116
Thr Gln Pro Tyr Thr Gln Asp Glu Ile Arg Asn Ile Leu Glu Ile Arg
                340                 345                 350

tgc caa gaa gag gat gtg gag atg aac gag gaa gcg aaa cag ctt ctg     1164
Cys Gln Glu Glu Asp Val Glu Met Asn Glu Glu Ala Lys Gln Leu Leu
            355                 360                 365

act ttg atc gga tgt aat acc tcg ctt agg tac gcg att cat cta atc     1212
Thr Leu Ile Gly Cys Asn Thr Ser Leu Arg Tyr Ala Ile His Leu Ile
        370                 375                 380

aat gca gct gcc cta gct tgc ctg aaa cgt aaa ggg aaa gtc gta gag     1260
Asn Ala Ala Ala Leu Ala Cys Leu Lys Arg Lys Gly Lys Val Val Glu
    385                 390                 395

att cag gac att gag aga gtt tat aga ttg ttt tta gac acc aag aga     1308
Ile Gln Asp Ile Glu Arg Val Tyr Arg Leu Phe Leu Asp Thr Lys Arg
400                 405                 410                 415

tcg atg cag tac ttg gtt gag cat gag agc gag tac ttg ttt agc gtg     1356
Ser Met Gln Tyr Leu Val Glu His Glu Ser Glu Tyr Leu Phe Ser Val
                420                 425                 430

cct ata aaa aac aca cag gag gct act gca gga gaagaaacag aacacgaggc   1409
Pro Ile Lys Asn Thr Gln Glu Ala Thr Ala Gly
            435                 440

catggaagtt tga                                                     1422
```

<210> 9

<211> 442

<212> PRT

<213> Arabidopsis thaliana

<400> 9

```
Ser His Ile Arg Gly Leu Gly Leu Asp Ser Val Leu Glu Pro Arg Ala
1               5                   10                  15

Val Ser Glu Gly Met Val Gly Gln Ile Lys Ala Arg Lys Ala Ala Gly
            20                  25                  30

Val Thr Leu Glu Leu Ile Arg Asp Gly Lys Ile Ser Gly Arg Ala Ile
        35                  40                  45

Leu Ile Ala Gly Gln Pro Gly Thr Gly Lys Ile Ala Ile Ala Met Gly
    50                  55                  60

Ile Ala Lys Ser Leu Gly Gln Glu Thr Pro Phe Thr Met Ile Ala Gly
65                  70                  75                  80

Ser Glu Ile Phe Ser Leu Glu Met Ser Lys Thr Glu Ala Leu Thr Gln
                85                  90                  95
```

```
Ala Phe Arg Lys Ala Ile Gly Val Arg Ile Lys Glu Glu Thr Asp Val
            100                 105                 110

Ile Glu Gly Glu Val Val Thr Ile Ser Ile Asp Arg Pro Ala Ser Ser
        115                 120                 125

Gly Gly Ser Val Lys Lys Thr Gly Lys Ile Thr Met Lys Thr Thr Asp
    130                 135                 140

Met Glu Ser Asn Phe Asp Leu Gly Trp Lys Leu Ile Glu Pro Leu Asp
145                 150                 155                 160

Lys Glu Lys Val Gln Ser Gly Asp Val Ile Val Leu Asp Arg Phe Cys
                165                 170                 175

Gly Lys Ile Thr Lys Leu Gly Arg Ser Phe Thr Arg Ser Arg Asp Phe
            180                 185                 190

Asp Val Met Gly Ser Lys Thr Lys Phe Val Gln Cys Pro Glu Gly Glu
        195                 200                 205

Leu Glu Lys Arg Lys Glu Val Leu His Ser Val Thr Leu His Glu Ile
    210                 215                 220

Asp Val Ile Asn Ser Arg Thr Gln Gly Tyr Leu Ala Leu Phe Thr Gly
225                 230                 235                 240

Asp Thr Gly Glu Ile Arg Ser Glu Thr Arg Glu Gln Ser Asp Thr Lys
                245                 250                 255

Val Ala Glu Trp Arg Glu Glu Gly Lys Ala Glu Ile Val Pro Gly Val
            260                 265                 270

Leu Phe Ile Asp Glu Val His Met Leu Asp Ile Glu Cys Phe Ser Phe
        275                 280                 285

Leu Asn Arg Ala Leu Glu Asn Asp Met Ser Pro Ile Leu Val Val Ala
    290                 295                 300

Thr Asn Arg Gly Met Thr Thr Ile Arg Gly Thr Asn Gln Ile Ser Ala
305                 310                 315                 320

His Gly Ile Pro Ile Asp Phe Leu Asp Arg Leu Leu Ile Ile Thr Thr
                325                 330                 335

Gln Pro Tyr Thr Gln Asp Glu Ile Arg Asn Ile Leu Glu Ile Arg Cys
            340                 345                 350

Gln Glu Glu Asp Val Glu Met Asn Glu Glu Ala Lys Gln Leu Leu Thr
```

355 360 365

Leu Ile Gly Cys Asn Thr Ser Leu Arg Tyr Ala Ile His Leu Ile Asn
370 375 380

Ala Ala Ala Leu Ala Cys Leu Lys Arg Lys Gly Lys Val Val Glu Ile
385 390 395 400

Gln Asp Ile Glu Arg Val Tyr Arg Leu Phe Leu Asp Thr Lys Arg Ser
405 410 415

Met Gln Tyr Leu Val Glu His Glu Ser Glu Tyr Leu Phe Ser Val Pro
420 425 430

Ile Lys Asn Thr Gln Glu Ala Thr Ala Gly
435 440

<210> 10

<211> 1591

<212> DNA

<213> Arabidopsis thaliana

<400> 10
ccacgcgtcc gtaagaagat tttgccagtg cgggaagcgg cggagattga gagattagcg 60
acgatggagg aaagcaaaca gaactatgac ctgacgccac taatagcgcc taacctggac 120
agacacttgg tgtttcctat attcgagttc cttcaagagc gtcagcttta ccctgatgag 180
cagatcctga agtctaaaat ccagcttttg aaccagacga acatggttga ttacgccatg 240
gatattcaca agagtctcta ccacactgaa gacgctcctc aagaaatggt ggagagaaga 300
acagaggttg tcgctaggct caaatctttg gaggaggctg ctgcaccact cgtgtctttt 360
cttttgaacc ctaacgctgt gcaggagcta agagctgaca agcagtacaa tctccaaatg 420
ctcaaggaac gctaccagat tggtccagac cagattgagg ctttgtacca gtacgccaag 480
tttcagtttg aatgtggcaa ctattctggt gctgctgatt atctttacca gtacaggacc 540
ctgtgctcta accttgagag gagtttgagt gccttgtggg gaaagctcgc atctgaaata 600
ttgatgcaaa actgggatat tgctcttgaa gagcttaacc gtctcaaaga gattattgac 660
tcaaagagtt tttcatcgcc gttaaaccag gtgcagaaca ggatttggtt gatgcattgg 720
ggtctgtata tctttttttaa ccatgataat ggaaggacac agatcattga tctttttaac 780
caagacaagt atctgaatgc catccaaact agtgctccac acttgctgcg ctacttggca 840
actgctttca ttgtcaacaa aaggagaaga ccacaattga aagaattcat taaggtcatt 900
cagcaagagc actactccta caaagatcca attatcgagt tcctggcatg tgtgtttgtc 960

aattatgact ttgatggggc tcaaaagaag atgaaagagt gtgaagaggt cattgtgaat 1020
gatccattcc ttggcaagcg agttgaggat ggaaactttt caactgtacc actgagagat 1080
gaatttcttg aaaatgcccg cctattcgtc tttgaaacct attgcaaaat tcatcaaagg 1140
attgacatgg gggtacttgc tgaaaaattg aatctgaact atgaggaggc cgagagatgg 1200
attgtgaacc taatccgcac ctcaaagctt gatgccaaga ttgattctga gtcaggaact 1260
gtaatcatgg agcctactca gcccaacgtg catgagcagt tgataaacca caccaaaggc 1320
ttatcaggac gaacatacaa gttagtgaat cagctcttgg aacacacaca ggcgcaagca 1380
actcgctagt caaaattttg ctgtggaagc ctttccttga taaaactcac cttcggttga 1440
ctggaattat tttctttttc ttgctctgag ttcacctttt actttgaaaa agattattat 1500
ggagttgttc tatttgaaat gttggatcca cagattggaa cattttccaa ccaaatcagc 1560
attttgtaaa aaaaaaaaaa aaaaaaaaaa a 1591

<210> 11
<211> 1773
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> CDS
<222> (1)..(1773)
<223>

<400> 11
atg gca gaa ctg cta ttt gag act tat gga gtg cct gca gtt gca ttt 48
Met Ala Glu Leu Leu Phe Glu Thr Tyr Gly Val Pro Ala Val Ala Phe
1 5 10 15

gga gtc gat gct gct ttc agc tac aaa tac aat caa cta cat gga att 96
Gly Val Asp Ala Ala Phe Ser Tyr Lys Tyr Asn Gln Leu His Gly Ile
20 25 30

tgt aaa aaa gat gga att gtt ctc tgt cct gga ttc acg aca aca cac 144
Cys Lys Lys Asp Gly Ile Val Leu Cys Pro Gly Phe Thr Thr Thr His
35 40 45

tcc att ccg ttt gtc gac gga gaa cct ata tat aaa gga tcc agc cga 192
Ser Ile Pro Phe Val Asp Gly Glu Pro Ile Tyr Lys Gly Ser Ser Arg
50 55 60

act aac att ggt gga tat cat gtc act gat tat tta aag cag ctt ctg 240
Thr Asn Ile Gly Gly Tyr His Val Thr Asp Tyr Leu Lys Gln Leu Leu
65 70 75 80

tca ctt aag tac cct ttt cat tcg tct agg ttt aca tgg gag aag gcc 288
Ser Leu Lys Tyr Pro Phe His Ser Ser Arg Phe Thr Trp Glu Lys Ala

```
                85                  90                  95

gaa gat ttg aaa ttg gaa cac tgt tat atc gca cct gac tat gct tcg      336
Glu Asp Leu Lys Leu Glu His Cys Tyr Ile Ala Pro Asp Tyr Ala Ser
            100                 105                 110

gaa att cgg tta ttc cag gaa gga aga aaa gaa gct gaa gag aaa aca      384
Glu Ile Arg Leu Phe Gln Glu Gly Arg Lys Glu Ala Glu Glu Lys Thr
        115                 120                 125

agt tat tgg cag ctt cca tgg ata cct cct ccc acc gaa gtt cct cca      432
Ser Tyr Trp Gln Leu Pro Trp Ile Pro Pro Pro Thr Glu Val Pro Pro
    130                 135                 140

tca gaa gaa gag att gca agg aag gca gct ata aga gaa aaa caa ggt      480
Ser Glu Glu Glu Ile Ala Arg Lys Ala Ala Ile Arg Glu Lys Gln Gly
145                 150                 155                 160

caa agg ctg cga gaa atg gct gaa gca aag aga gtg tcc aag att aat      528
Gln Arg Leu Arg Glu Met Ala Glu Ala Lys Arg Val Ser Lys Ile Asn
                165                 170                 175

gac atg gag aat caa ctg att agc ttg cgt ttc ctt ttg aag caa gtt      576
Asp Met Glu Asn Gln Leu Ile Ser Leu Arg Phe Leu Leu Lys Gln Val
            180                 185                 190

gac cag gtt gaa gag gat gat att cca acc ttt ttg tca gac acc ggt      624
Asp Gln Val Glu Glu Asp Asp Ile Pro Thr Phe Leu Ser Asp Thr Gly
        195                 200                 205

tac gcg tcc agg caa gag cta gag tct act att acg aaa gtg aca cag      672
Tyr Ala Ser Arg Gln Glu Leu Glu Ser Thr Ile Thr Lys Val Thr Gln
    210                 215                 220

tcg ctt aga aaa gca agg ggt gag ccg aag aat gaa cca gct gag tat      720
Ser Leu Arg Lys Ala Arg Gly Glu Pro Lys Asn Glu Pro Ala Glu Tyr
225                 230                 235                 240

gaa gaa aac cct gat tct ctt aat aat gaa aag tat cca ctt atg aat      768
Glu Glu Asn Pro Asp Ser Leu Asn Asn Glu Lys Tyr Pro Leu Met Asn
                245                 250                 255

gtc ccc gat gat att ctt act cct gag cag ctt aag gac aag aag agg      816
Val Pro Asp Asp Ile Leu Thr Pro Glu Gln Leu Lys Asp Lys Lys Arg
            260                 265                 270

caa atg ttt ctt aaa aca aca gca gag ggc cgg cta cga gct aga caa      864
Gln Met Phe Leu Lys Thr Thr Ala Glu Gly Arg Leu Arg Ala Arg Gln
        275                 280                 285

aag cgt aat gag gag gaa ctc gaa aaa gag aaa aga aat caa tta gag      912
Lys Arg Asn Glu Glu Glu Leu Glu Lys Glu Lys Arg Asn Gln Leu Glu
    290                 295                 300

gag gaa aga cgt cgt gag aac cca gag tct tac tta gag gag ttg caa      960
Glu Glu Arg Arg Arg Glu Asn Pro Glu Ser Tyr Leu Glu Glu Leu Gln
305                 310                 315                 320

gct cag tac aag gaa gtg ttg gag aga gtt gag cag aag aag cgt ctg     1008
Ala Gln Tyr Lys Glu Val Leu Glu Arg Val Glu Gln Lys Lys Arg Leu
                325                 330                 335

aaa aca aac ggg tcc agt aac ggg aat aac aag tct gga ggt att ggg     1056
Lys Thr Asn Gly Ser Ser Asn Gly Asn Asn Lys Ser Gly Gly Ile Gly
            340                 345                 350
```

```
cga ggc gag cga ctc agt gct gca cag agg gag aga atg cgt ctg ctg      1104
Arg Gly Glu Arg Leu Ser Ala Ala Gln Arg Glu Arg Met Arg Leu Leu
        355                 360                 365

acg aca gca gcc ttt gat aga ggg aaa ggc gag gat acg ttt ggt tct      1152
Thr Thr Ala Ala Phe Asp Arg Gly Lys Gly Glu Asp Thr Phe Gly Ser
    370                 375                 380

aga gat gaa gat tgg cag ctc tac aaa ctt atg agc aag gat aat gac      1200
Arg Asp Glu Asp Trp Gln Leu Tyr Lys Leu Met Ser Lys Asp Asn Asp
385                 390                 395                 400

gat gat gac gaa caa cct gat tca gac gag gca gag ttg gct cgt tta      1248
Asp Asp Asp Glu Gln Pro Asp Ser Asp Glu Ala Glu Leu Ala Arg Leu
                405                 410                 415

tca tct aga ctt cag gaa att gat cca aca ttt gtg cag aaa gta gaa      1296
Ser Ser Arg Leu Gln Glu Ile Asp Pro Thr Phe Val Gln Lys Val Glu
            420                 425                 430

gga gaa ttg agt cag aca tca ggg gag gtg cca cgc gta cgc cca tta      1344
Gly Glu Leu Ser Gln Thr Ser Gly Glu Val Pro Arg Val Arg Pro Leu
        435                 440                 445

aca gag gaa gac tac aag ata gtg att ggt ata gaa aga ttc cgt tgc      1392
Thr Glu Glu Asp Tyr Lys Ile Val Ile Gly Ile Glu Arg Phe Arg Cys
    450                 455                 460

cca gag atc ctg ttc cat cca aac ctt att ggt att gac caa gta ggc      1440
Pro Glu Ile Leu Phe His Pro Asn Leu Ile Gly Ile Asp Gln Val Gly
465                 470                 475                 480

tta gac gag atg gct ggc aca tca atc aga agg cta ccg cac gac gag      1488
Leu Asp Glu Met Ala Gly Thr Ser Ile Arg Arg Leu Pro His Asp Glu
                485                 490                 495

aaa gag tta gag gag agg cta acg agc tcg ata cta atg acg ggc ggg      1536
Lys Glu Leu Glu Glu Arg Leu Thr Ser Ser Ile Leu Met Thr Gly Gly
            500                 505                 510

tgt agc ctt ctt cca ggg atg aac gag cgg ttg gaa tgt ggg att agg      1584
Cys Ser Leu Leu Pro Gly Met Asn Glu Arg Leu Glu Cys Gly Ile Arg
        515                 520                 525

atg ata aga cct tgc gga tca ccc att aac gtg gtt aga gct atg gat      1632
Met Ile Arg Pro Cys Gly Ser Pro Ile Asn Val Val Arg Ala Met Asp
    530                 535                 540

cca gtg ctg gat gca tgg cga gga gca tct gca ttt gct gct aat ttg      1680
Pro Val Leu Asp Ala Trp Arg Gly Ala Ser Ala Phe Ala Ala Asn Leu
545                 550                 555                 560

aac ttc ttg ggg aat gcc ttt act aag atg gat tac gac gag aaa ggt      1728
Asn Phe Leu Gly Asn Ala Phe Thr Lys Met Asp Tyr Asp Glu Lys Gly
                565                 570                 575

gaa gat tgg ctt aga aat tat caa att cga tac aac tat ttg tga          1773
Glu Asp Trp Leu Arg Asn Tyr Gln Ile Arg Tyr Asn Tyr Leu
            580                 585                 590
```

<210> 12

<211> 590

<212> PRT

<213> Arabidopsis thaliana

<400> 12

```
Met Ala Glu Leu Leu Phe Glu Thr Tyr Gly Val Pro Ala Val Ala Phe
1               5                  10                  15

Gly Val Asp Ala Ala Phe Ser Tyr Lys Tyr Asn Gln Leu His Gly Ile
            20                  25                  30

Cys Lys Lys Asp Gly Ile Val Leu Cys Pro Gly Phe Thr Thr Thr His
        35                  40                  45

Ser Ile Pro Phe Val Asp Gly Glu Pro Ile Tyr Lys Gly Ser Ser Arg
    50                  55                  60

Thr Asn Ile Gly Gly Tyr His Val Thr Asp Tyr Leu Lys Gln Leu Leu
65                  70                  75                  80

Ser Leu Lys Tyr Pro Phe His Ser Ser Arg Phe Thr Trp Glu Lys Ala
                85                  90                  95

Glu Asp Leu Lys Leu Glu His Cys Tyr Ile Ala Pro Asp Tyr Ala Ser
            100                 105                 110

Glu Ile Arg Leu Phe Gln Glu Gly Arg Lys Glu Ala Glu Glu Lys Thr
        115                 120                 125

Ser Tyr Trp Gln Leu Pro Trp Ile Pro Pro Pro Thr Glu Val Pro Pro
    130                 135                 140

Ser Glu Glu Glu Ile Ala Arg Lys Ala Ala Ile Arg Glu Lys Gln Gly
145                 150                 155                 160

Gln Arg Leu Arg Glu Met Ala Glu Ala Lys Arg Val Ser Lys Ile Asn
                165                 170                 175

Asp Met Glu Asn Gln Leu Ile Ser Leu Arg Phe Leu Leu Lys Gln Val
            180                 185                 190

Asp Gln Val Glu Glu Asp Asp Ile Pro Thr Phe Leu Ser Asp Thr Gly
        195                 200                 205

Tyr Ala Ser Arg Gln Glu Leu Glu Ser Thr Ile Thr Lys Val Thr Gln
    210                 215                 220

Ser Leu Arg Lys Ala Arg Gly Glu Pro Lys Asn Glu Pro Ala Glu Tyr
225                 230                 235                 240
```

```
Glu Glu Asn Pro Asp Ser Leu Asn Asn Glu Lys Tyr Pro Leu Met Asn
                245                 250                 255

Val Pro Asp Asp Ile Leu Thr Pro Glu Gln Leu Lys Asp Lys Lys Arg
            260                 265                 270

Gln Met Phe Leu Lys Thr Thr Ala Glu Gly Arg Leu Arg Ala Arg Gln
        275                 280                 285

Lys Arg Asn Glu Glu Glu Leu Glu Lys Glu Lys Arg Asn Gln Leu Glu
    290                 295                 300

Glu Glu Arg Arg Arg Glu Asn Pro Glu Ser Tyr Leu Glu Glu Leu Gln
305                 310                 315                 320

Ala Gln Tyr Lys Glu Val Leu Glu Arg Val Glu Gln Lys Lys Arg Leu
                325                 330                 335

Lys Thr Asn Gly Ser Ser Asn Gly Asn Asn Lys Ser Gly Gly Ile Gly
            340                 345                 350

Arg Gly Glu Arg Leu Ser Ala Ala Gln Arg Glu Arg Met Arg Leu Leu
        355                 360                 365

Thr Thr Ala Ala Phe Asp Arg Gly Lys Gly Glu Asp Thr Phe Gly Ser
    370                 375                 380

Arg Asp Glu Asp Trp Gln Leu Tyr Lys Leu Met Ser Lys Asp Asn Asp
385                 390                 395                 400

Asp Asp Asp Glu Gln Pro Asp Ser Asp Glu Ala Glu Leu Ala Arg Leu
                405                 410                 415

Ser Ser Arg Leu Gln Glu Ile Asp Pro Thr Phe Val Gln Lys Val Glu
            420                 425                 430

Gly Glu Leu Ser Gln Thr Ser Gly Glu Val Pro Arg Val Arg Pro Leu
        435                 440                 445

Thr Glu Glu Asp Tyr Lys Ile Val Ile Gly Ile Glu Arg Phe Arg Cys
    450                 455                 460

Pro Glu Ile Leu Phe His Pro Asn Leu Ile Gly Ile Asp Gln Val Gly
465                 470                 475                 480

Leu Asp Glu Met Ala Gly Thr Ser Ile Arg Arg Leu Pro His Asp Glu
                485                 490                 495
```

Lys Glu Leu Glu Glu Arg Leu Thr Ser Ser Ile Leu Met Thr Gly Gly
500 505 510

Cys Ser Leu Leu Pro Gly Met Asn Glu Arg Leu Glu Cys Gly Ile Arg
515 520 525

Met Ile Arg Pro Cys Gly Ser Pro Ile Asn Val Val Arg Ala Met Asp
530 535 540

Pro Val Leu Asp Ala Trp Arg Gly Ala Ser Ala Phe Ala Ala Asn Leu
545 550 555 560

Asn Phe Leu Gly Asn Ala Phe Thr Lys Met Asp Tyr Asp Glu Lys Gly
565 570 575

Glu Asp Trp Leu Arg Asn Tyr Gln Ile Arg Tyr Asn Tyr Leu
580 585 590

<210> 13

<211> 1416

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(1416)

<223>

<400> 13

atg atc ctg aag aaa gta tgg gga tcg gtg tgg aac cga tcg aat agt 48
Met Ile Leu Lys Lys Val Trp Gly Ser Val Trp Asn Arg Ser Asn Ser
1 5 10 15

ggc aag gat ttg gtg aat cat cag aga gcg atc gat gtt cct cct ctg 96
Gly Lys Asp Leu Val Asn His Gln Arg Ala Ile Asp Val Pro Pro Leu
20 25 30

tta ttg tct tct tca tcg tct ctt ggt gcg ttt gat cag cta ccg atg 144
Leu Leu Ser Ser Ser Ser Ser Leu Gly Ala Phe Asp Gln Leu Pro Met
35 40 45

gat att cta gtc cag ata ctg atg atg atg gag cca aaa gat gct gtg 192
Asp Ile Leu Val Gln Ile Leu Met Met Met Glu Pro Lys Asp Ala Val
50 55 60

aaa ttg ggc tta acg tgc aaa gcc tgg aaa tgc gta gct agt ggt aat 240
Lys Leu Gly Leu Thr Cys Lys Ala Trp Lys Cys Val Ala Ser Gly Asn
65 70 75 80

cgt ctc tgg ata ttt tat ctc cag tgt tct caa gag cca tgg gac tcc 288

```
Arg Leu Trp Ile Phe Tyr Leu Gln Cys Ser Gln Glu Pro Trp Asp Ser
                85                  90                  95

att ttc ttc gct gaa act agt ttg cgt tct ggt tat cct ctc cga atg      336
Ile Phe Phe Ala Glu Thr Ser Leu Arg Ser Gly Tyr Pro Leu Arg Met
            100                 105                 110

att tct agt caa tca gga gag ttg tcg ttt atg cac att tat agt cag      384
Ile Ser Ser Gln Ser Gly Glu Leu Ser Phe Met His Ile Tyr Ser Gln
        115                 120                 125

agg gca caa gtt cct ggt tct atc att att gat ggt ggt tct gga tat      432
Arg Ala Gln Val Pro Gly Ser Ile Ile Ile Asp Gly Gly Ser Gly Tyr
    130                 135                 140

tgt aag ttt ggt tgg agc aag tat gcg tct cct tct gga cgt tct gct      480
Cys Lys Phe Gly Trp Ser Lys Tyr Ala Ser Pro Ser Gly Arg Ser Ala
145                 150                 155                 160

act ttt ttg gaa ttt ggt aac att gag tca ccg att tat gct aga ctt      528
Thr Phe Leu Glu Phe Gly Asn Ile Glu Ser Pro Ile Tyr Ala Arg Leu
                165                 170                 175

caa cag ttc ttt gca acc att ttc acc agg atg cag gta aag ccc tct      576
Gln Gln Phe Phe Ala Thr Ile Phe Thr Arg Met Gln Val Lys Pro Ser
            180                 185                 190

atg cag cca ata gtg gta tca cta cct ctc tgc cat ttt gat gat act      624
Met Gln Pro Ile Val Val Ser Leu Pro Leu Cys His Phe Asp Asp Thr
        195                 200                 205

gaa tca gcc aag gca tca agg cgg caa ctt aag act gct att ttc aat      672
Glu Ser Ala Lys Ala Ser Arg Arg Gln Leu Lys Thr Ala Ile Phe Asn
    210                 215                 220

gtc ttg ttt gac atg aac gtc cct gca gtg tgt gca gtt aat cag gct      720
Val Leu Phe Asp Met Asn Val Pro Ala Val Cys Ala Val Asn Gln Ala
225                 230                 235                 240

gtg tta gct cta tat gca gca cgg cgg aca tct gga att gtt gtc aac      768
Val Leu Ala Leu Tyr Ala Ala Arg Arg Thr Ser Gly Ile Val Val Asn
                245                 250                 255

att ggt ttc caa gtc atc acc att ctt ccg att tta cat ggt aag gtg      816
Ile Gly Phe Gln Val Ile Thr Ile Leu Pro Ile Leu His Gly Lys Val
            260                 265                 270

atg cgc cag gta ggt gta gaa gtc att ggt ttt gga gca ttg aaa ctc      864
Met Arg Gln Val Gly Val Glu Val Ile Gly Phe Gly Ala Leu Lys Leu
        275                 280                 285

acg ggc ttc ctt aag gag aag atg caa gaa aac aac att tcc ttt caa      912
Thr Gly Phe Leu Lys Glu Lys Met Gln Glu Asn Asn Ile Ser Phe Gln
    290                 295                 300

tca ctc tac act gtt cgt act ctt aaa gag aaa ctg tgt tat gtg gct      960
Ser Leu Tyr Thr Val Arg Thr Leu Lys Glu Lys Leu Cys Tyr Val Ala
305                 310                 315                 320

ctt gat tat aaa gct gaa ctt tca aaa gac aca caa gct tca gtg gaa     1008
Leu Asp Tyr Lys Ala Glu Leu Ser Lys Asp Thr Gln Ala Ser Val Glu
                325                 330                 335

gtt tca ggt gaa gga tgg ttt act tta tca aaa gag cgt ttc caa aca     1056
Val Ser Gly Glu Gly Trp Phe Thr Leu Ser Lys Glu Arg Phe Gln Thr
            340                 345                 350
```

```
ggg gag ata tta ttc caa cca cgt ctc gct gga atg cgt gca atg agt     1104
Gly Glu Ile Leu Phe Gln Pro Arg Leu Ala Gly Met Arg Ala Met Ser
        355                 360                 365

ctg cac cag gcc gtc tcg ctc tgt atg gac cac tgt gat gca gca gga     1152
Leu His Gln Ala Val Ser Leu Cys Met Asp His Cys Asp Ala Ala Gly
    370                 375                 380

ctt aca ggt gac gat agt tgg ttc aag act gta gta cta act ggg gga     1200
Leu Thr Gly Asp Asp Ser Trp Phe Lys Thr Val Val Leu Thr Gly Gly
385                 390                 395                 400

agc gcg tgt ttg cct gga ctc tca gag agg cta gag aga gaa ctg caa     1248
Ser Ala Cys Leu Pro Gly Leu Ser Glu Arg Leu Glu Arg Glu Leu Gln
                405                 410                 415

gat cac ctt cct tca tct ata agt aac gga atc aga gta ata cct cct     1296
Asp His Leu Pro Ser Ser Ile Ser Asn Gly Ile Arg Val Ile Pro Pro
            420                 425                 430

cct tac ggc gtg gac aca tca tgg cat ggg gca aag ctt att agt aat     1344
Pro Tyr Gly Val Asp Thr Ser Trp His Gly Ala Lys Leu Ile Ser Asn
        435                 440                 445

ttg agc atc ttt cct ggt cct tgg tgt atc aca agg aag cag ttc cgt     1392
Leu Ser Ile Phe Pro Gly Pro Trp Cys Ile Thr Arg Lys Gln Phe Arg
    450                 455                 460

cgc aag tca aga ctc atg tgg tga                                    1416
Arg Lys Ser Arg Leu Met Trp
465                 470
```

<210> 14

<211> 471

<212> PRT

<213> Arabidopsis thaliana

<400> 14

```
Met Ile Leu Lys Lys Val Trp Gly Ser Val Trp Asn Arg Ser Asn Ser
1               5                   10                  15

Gly Lys Asp Leu Val Asn His Gln Arg Ala Ile Asp Val Pro Pro Leu
            20                  25                  30

Leu Leu Ser Ser Ser Ser Ser Leu Gly Ala Phe Asp Gln Leu Pro Met
        35                  40                  45

Asp Ile Leu Val Gln Ile Leu Met Met Met Glu Pro Lys Asp Ala Val
    50                  55                  60

Lys Leu Gly Leu Thr Cys Lys Ala Trp Lys Cys Val Ala Ser Gly Asn
65                  70                  75                  80
```

```
Arg Leu Trp Ile Phe Tyr Leu Gln Cys Ser Gln Glu Pro Trp Asp Ser
                85                  90                  95

Ile Phe Phe Ala Glu Thr Ser Leu Arg Ser Gly Tyr Pro Leu Arg Met
            100                 105                 110

Ile Ser Ser Gln Ser Gly Glu Leu Ser Phe Met His Ile Tyr Ser Gln
        115                 120                 125

Arg Ala Gln Val Pro Gly Ser Ile Ile Ile Asp Gly Gly Ser Gly Tyr
    130                 135                 140

Cys Lys Phe Gly Trp Ser Lys Tyr Ala Ser Pro Ser Gly Arg Ser Ala
145                 150                 155                 160

Thr Phe Leu Glu Phe Gly Asn Ile Glu Ser Pro Ile Tyr Ala Arg Leu
                165                 170                 175

Gln Gln Phe Phe Ala Thr Ile Phe Thr Arg Met Gln Val Lys Pro Ser
            180                 185                 190

Met Gln Pro Ile Val Val Ser Leu Pro Leu Cys His Phe Asp Asp Thr
        195                 200                 205

Glu Ser Ala Lys Ala Ser Arg Arg Gln Leu Lys Thr Ala Ile Phe Asn
    210                 215                 220

Val Leu Phe Asp Met Asn Val Pro Ala Val Cys Ala Val Asn Gln Ala
225                 230                 235                 240

Val Leu Ala Leu Tyr Ala Ala Arg Arg Thr Ser Gly Ile Val Val Asn
                245                 250                 255

Ile Gly Phe Gln Val Ile Thr Ile Leu Pro Ile Leu His Gly Lys Val
            260                 265                 270

Met Arg Gln Val Gly Val Glu Val Ile Gly Phe Gly Ala Leu Lys Leu
        275                 280                 285

Thr Gly Phe Leu Lys Glu Lys Met Gln Glu Asn Asn Ile Ser Phe Gln
    290                 295                 300

Ser Leu Tyr Thr Val Arg Thr Leu Lys Glu Lys Leu Cys Tyr Val Ala
305                 310                 315                 320

Leu Asp Tyr Lys Ala Glu Leu Ser Lys Asp Thr Gln Ala Ser Val Glu
                325                 330                 335

Val Ser Gly Glu Gly Trp Phe Thr Leu Ser Lys Glu Arg Phe Gln Thr
            340                 345                 350
```

Gly Glu Ile Leu Phe Gln Pro Arg Leu Ala Gly Met Arg Ala Met Ser
355 360 365

Leu His Gln Ala Val Ser Leu Cys Met Asp His Cys Asp Ala Ala Gly
370 375 380

Leu Thr Gly Asp Asp Ser Trp Phe Lys Thr Val Val Leu Thr Gly Gly
385 390 395 400

Ser Ala Cys Leu Pro Gly Leu Ser Glu Arg Leu Glu Arg Glu Leu Gln
405 410 415

Asp His Leu Pro Ser Ser Ile Ser Asn Gly Ile Arg Val Ile Pro Pro
420 425 430

Pro Tyr Gly Val Asp Thr Ser Trp His Gly Ala Lys Leu Ile Ser Asn
435 440 445

Leu Ser Ile Phe Pro Gly Pro Trp Cys Ile Thr Arg Lys Gln Phe Arg
450 455 460

Arg Lys Ser Arg Leu Met Trp
465 470

<210> 15

<211> 1092

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(1092)

<223>

<400> 15

atg gaa gca cta gtt gtt gat gct ggc tct aag ttc ctg aaa gca gga 48
Met Glu Ala Leu Val Val Asp Ala Gly Ser Lys Phe Leu Lys Ala Gly
1 5 10 15

gca gca att cct gac cag tct cct gca atg ata att ccc tct caa atg 96
Ala Ala Ile Pro Asp Gln Ser Pro Ala Met Ile Ile Pro Ser Gln Met
20 25 30

aaa cga atg gtt gat gat ggg tct tct tca gct gat aac ccc acc act 144
Lys Arg Met Val Asp Asp Gly Ser Ser Ser Ala Asp Asn Pro Thr Thr
35 40 45

```
gtc ttt gag gat gtc act ctt gat cct att gaa agg ggt ttg att aga      192
Val Phe Glu Asp Val Thr Leu Asp Pro Ile Glu Arg Gly Leu Ile Arg
    50                  55                  60

gat tgg gat gct atg gaa gat ctg ttg cgt tat gtt gtc tac act ggg      240
Asp Trp Asp Ala Met Glu Asp Leu Leu Arg Tyr Val Val Tyr Thr Gly
65                  70                  75                  80

ctt gga tgg gaa gag gga aac gaa ggc aat ata ctt ttt aca gat cca      288
Leu Gly Trp Glu Glu Gly Asn Glu Gly Asn Ile Leu Phe Thr Asp Pro
                85                  90                  95

ctt tgt act cct aag gct att agg gag caa ttg gtg cag ttg atg ttt      336
Leu Cys Thr Pro Lys Ala Ile Arg Glu Gln Leu Val Gln Leu Met Phe
            100                 105                 110

gaa aca ttc aat gtc tct gga ttt tat gca tct gag caa gca gtg ttg      384
Glu Thr Phe Asn Val Ser Gly Phe Tyr Ala Ser Glu Gln Ala Val Leu
        115                 120                 125

tcc ctt tat gct gtt gga cgc atc tcc ggt tgc act gtt gat att ggt      432
Ser Leu Tyr Ala Val Gly Arg Ile Ser Gly Cys Thr Val Asp Ile Gly
    130                 135                 140

cat ggg aag ata gat att gcc cca gtt ctt gaa ggt gca gta caa cac      480
His Gly Lys Ile Asp Ile Ala Pro Val Leu Glu Gly Ala Val Gln His
145                 150                 155                 160

att gcc tcg aaa cgg ttt gag cta ggt gga acc gag cta act aaa tta      528
Ile Ala Ser Lys Arg Phe Glu Leu Gly Gly Thr Glu Leu Thr Lys Leu
                165                 170                 175

ttt gcc caa gag ctt gga aaa acc aac ccg tcg atg aat ctc agc atg      576
Phe Ala Gln Glu Leu Gly Lys Thr Asn Pro Ser Met Asn Leu Ser Met
            180                 185                 190

tct gat gtt gaa aaa ctc aag gag cag tat gca aac tgt gcc gag gac      624
Ser Asp Val Glu Lys Leu Lys Glu Gln Tyr Ala Asn Cys Ala Glu Asp
        195                 200                 205

gaa att gct tac aaa aaa acc caa aac tgt gaa atc gag cag cat act      672
Glu Ile Ala Tyr Lys Lys Thr Gln Asn Cys Glu Ile Glu Gln His Thr
    210                 215                 220

ctt cct gat gga cag gtg ata agc atc ggg cga gag aga tac tcg gtt      720
Leu Pro Asp Gly Gln Val Ile Ser Ile Gly Arg Glu Arg Tyr Ser Val
225                 230                 235                 240

gga gaa gct ctg ttt cag cca tca ata ctg gga ctg gag gag cat gga      768
Gly Glu Ala Leu Phe Gln Pro Ser Ile Leu Gly Leu Glu Glu His Gly
                245                 250                 255

atc gtt gag cag ctt gtc cgg att atc tcc aca gtg tca tct gag aac      816
Ile Val Glu Gln Leu Val Arg Ile Ile Ser Thr Val Ser Ser Glu Asn
            260                 265                 270

cat agg cag ctc ttg gag aac act gta ctt tgt ggt ggt aca acc tcc      864
His Arg Gln Leu Leu Glu Asn Thr Val Leu Cys Gly Gly Thr Thr Ser
        275                 280                 285

atg aca gga ttc gaa agt aga ttc cag aaa gaa gca aac ttg tgc tca      912
Met Thr Gly Phe Glu Ser Arg Phe Gln Lys Glu Ala Asn Leu Cys Ser
    290                 295                 300

tct gcc att agg cca aca ctg gtg aaa ccg cca gaa tat atg ccg gag      960
```

```
Ser Ala Ile Arg Pro Thr Leu Val Lys Pro Pro Glu Tyr Met Pro Glu
305                 310                 315                 320

aat ttg ggg atg tat tcg gct tgg gtt gga gga gcc ata cta gct aaa      1008
Asn Leu Gly Met Tyr Ser Ala Trp Val Gly Gly Ala Ile Leu Ala Lys
                325                 330                 335

gtg gtg ttt ccg cag aat cag cac gtt act aaa gca gat tat gac gag      1056
Val Val Phe Pro Gln Asn Gln His Val Thr Lys Ala Asp Tyr Asp Glu
            340                 345                 350

act gga cca tca gtg gtt cac agg aaa tgt ttc tga                     1092
Thr Gly Pro Ser Val Val His Arg Lys Cys Phe
        355                 360
```

<210> 16

<211> 363

<212> PRT

<213> Arabidopsis thaliana

<400> 16

```
Met Glu Ala Leu Val Val Asp Ala Gly Ser Lys Phe Leu Lys Ala Gly
1               5                   10                  15

Ala Ala Ile Pro Asp Gln Ser Pro Ala Met Ile Ile Pro Ser Gln Met
            20                  25                  30

Lys Arg Met Val Asp Asp Gly Ser Ser Ser Ala Asp Asn Pro Thr Thr
        35                  40                  45

Val Phe Glu Asp Val Thr Leu Asp Pro Ile Glu Arg Gly Leu Ile Arg
    50                  55                  60

Asp Trp Asp Ala Met Glu Asp Leu Leu Arg Tyr Val Val Tyr Thr Gly
65                  70                  75                  80

Leu Gly Trp Glu Glu Gly Asn Glu Gly Asn Ile Leu Phe Thr Asp Pro
                85                  90                  95

Leu Cys Thr Pro Lys Ala Ile Arg Glu Gln Leu Val Gln Leu Met Phe
            100                 105                 110

Glu Thr Phe Asn Val Ser Gly Phe Tyr Ala Ser Glu Gln Ala Val Leu
        115                 120                 125

Ser Leu Tyr Ala Val Gly Arg Ile Ser Gly Cys Thr Val Asp Ile Gly
    130                 135                 140

His Gly Lys Ile Asp Ile Ala Pro Val Leu Glu Gly Ala Val Gln His
145                 150                 155                 160
```

Ile Ala Ser Lys Arg Phe Glu Leu Gly Gly Thr Glu Leu Thr Lys Leu
165 170 175

Phe Ala Gln Glu Leu Gly Lys Thr Asn Pro Ser Met Asn Leu Ser Met
180 185 190

Ser Asp Val Glu Lys Leu Lys Glu Gln Tyr Ala Asn Cys Ala Glu Asp
195 200 205

Glu Ile Ala Tyr Lys Lys Thr Gln Asn Cys Glu Ile Glu Gln His Thr
210 215 220

Leu Pro Asp Gly Gln Val Ile Ser Ile Gly Arg Glu Arg Tyr Ser Val
225 230 235 240

Gly Glu Ala Leu Phe Gln Pro Ser Ile Leu Gly Leu Glu Glu His Gly
245 250 255

Ile Val Glu Gln Leu Val Arg Ile Ile Ser Thr Val Ser Ser Glu Asn
260 265 270

His Arg Gln Leu Leu Glu Asn Thr Val Leu Cys Gly Gly Thr Thr Ser
275 280 285

Met Thr Gly Phe Glu Ser Arg Phe Gln Lys Glu Ala Asn Leu Cys Ser
290 295 300

Ser Ala Ile Arg Pro Thr Leu Val Lys Pro Pro Glu Tyr Met Pro Glu
305 310 315 320

Asn Leu Gly Met Tyr Ser Ala Trp Val Gly Gly Ala Ile Leu Ala Lys
325 330 335

Val Val Phe Pro Gln Asn Gln His Val Thr Lys Ala Asp Tyr Asp Glu
340 345 350

Thr Gly Pro Ser Val Val His Arg Lys Cys Phe
355 360

<210> 17

<211> 1791

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(1791)

<223>

<400> 17

```
atg gca aaa cca aaa tcc aat tct cat ctc tcg tgg cag gat tac ttg       48
Met Ala Lys Pro Lys Ser Asn Ser His Leu Ser Trp Gln Asp Tyr Leu
1               5                  10                  15

aaa act gtt gcc ccc act cag att ttg tct gaa cga gga gct aat ctc       96
Lys Thr Val Ala Pro Thr Gln Ile Leu Ser Glu Arg Gly Ala Asn Leu
            20                  25                  30

gtt gtc atc aac tta ggg tct gct aat gta aga gta gga ttg gct atg      144
Val Val Ile Asn Leu Gly Ser Ala Asn Val Arg Val Gly Leu Ala Met
        35                  40                  45

gat gag aaa cct ttc aat gtt cct aat tgt att gca aga tac att aca      192
Asp Glu Lys Pro Phe Asn Val Pro Asn Cys Ile Ala Arg Tyr Ile Thr
    50                  55                  60

caa tct ggg aaa cca act gtt gtt gat cag atg ctt aac act gag gtc      240
Gln Ser Gly Lys Pro Thr Val Val Asp Gln Met Leu Asn Thr Glu Val
65                  70                  75                  80

act act aat caa cat gtt gac cgt gag aga gct tat aat tct gcg gca      288
Thr Thr Asn Gln His Val Asp Arg Glu Arg Ala Tyr Asn Ser Ala Ala
                85                  90                  95

tcg ctg ttg aag ata ctg ttc ctg gat gaa agt tct tct tct ggc tct      336
Ser Leu Leu Lys Ile Leu Phe Leu Asp Glu Ser Ser Ser Ser Gly Ser
            100                 105                 110

gca tct cgc aag atg ggg agg atc gac gga tac aat caa gcg agt acg      384
Ala Ser Arg Lys Met Gly Arg Ile Asp Gly Tyr Asn Gln Ala Ser Thr
        115                 120                 125

ata aag aaa gat tca gtc ttt act tgg act gat gtg tat gag gat gag      432
Ile Lys Lys Asp Ser Val Phe Thr Trp Thr Asp Val Tyr Glu Asp Glu
    130                 135                 140

aaa att tct ctt gcc tca cca gca gaa act tct cct gat aaa ggt gat      480
Lys Ile Ser Leu Ala Ser Pro Ala Glu Thr Ser Pro Asp Lys Gly Asp
145                 150                 155                 160

gct agt gca tct gaa gct gtt cct gat gtt act gac tct aaa gat act      528
Ala Ser Ala Ser Glu Ala Val Pro Asp Val Thr Asp Ser Lys Asp Thr
                165                 170                 175

agt gag agt aaa cgc aag tat agg aaa atg att ttt ggt gaa gaa gct      576
Ser Glu Ser Lys Arg Lys Tyr Arg Lys Met Ile Phe Gly Glu Glu Ala
            180                 185                 190

ttg aaa ata tcg cca aaa gag cca tat tgt tta tac cat cct att cgg      624
Leu Lys Ile Ser Pro Lys Glu Pro Tyr Cys Leu Tyr His Pro Ile Arg
        195                 200                 205

aga ggt cac ttc aat gtt tcg cca cat tat tca gcg caa cgg gtt tgt      672
Arg Gly His Phe Asn Val Ser Pro His Tyr Ser Ala Gln Arg Val Cys
    210                 215                 220

gag gac tta act gct atc ttg gac tgg att tta tta gag aaa ctt cat      720
```

```
Glu Asp Leu Thr Ala Ile Leu Asp Trp Ile Leu Leu Glu Lys Leu His
225                 230                 235                 240

ata act cac aag gag aga ttc tct ttt cat gct gtt att gtt gtc cca      768
Ile Thr His Lys Glu Arg Phe Ser Phe His Ala Val Ile Val Val Pro
                245                 250                 255

gaa aca ttt gac acc cgc gaa ata aag gaa atg cta act att gtg ttg      816
Glu Thr Phe Asp Thr Arg Glu Ile Lys Glu Met Leu Thr Ile Val Leu
            260                 265                 270

gga gag cta tac ttt aac tca gca gtt gtc cac caa gaa ggt cta tcg      864
Gly Glu Leu Tyr Phe Asn Ser Ala Val Val His Gln Glu Gly Leu Ser
        275                 280                 285

gcc gtt ttt ggg aat ggt ttg aca aca gct tgt att gtg aat ata gga      912
Ala Val Phe Gly Asn Gly Leu Thr Thr Ala Cys Ile Val Asn Ile Gly
    290                 295                 300

gcc cag aca agt aca gta gtt tgt gtc gag gat ggg gtc tca ttg cca      960
Ala Gln Thr Ser Thr Val Val Cys Val Glu Asp Gly Val Ser Leu Pro
305                 310                 315                 320

aat act gaa aag att tta cct ttt gga gga gag gat ata tgt aga tgc     1008
Asn Thr Glu Lys Ile Leu Pro Phe Gly Gly Glu Asp Ile Cys Arg Cys
                325                 330                 335

ctt cta tgg att cag agg cat tac caa aag tgg cca caa atc aac aca     1056
Leu Leu Trp Ile Gln Arg His Tyr Gln Lys Trp Pro Gln Ile Asn Thr
            340                 345                 350

gat gtt ttg gca aag cca atc gat atg ctg atg ctt aat caa ctt aag     1104
Asp Val Leu Ala Lys Pro Ile Asp Met Leu Met Leu Asn Gln Leu Lys
        355                 360                 365

gag tca ttt tgt gaa att aga gca gga gaa ctt gaa act gtt gca acg     1152
Glu Ser Phe Cys Glu Ile Arg Ala Gly Glu Leu Glu Thr Val Ala Thr
    370                 375                 380

gtt cat tct tat gag gaa ggc atg cca gct gtg cct cac aag aca aat     1200
Val His Ser Tyr Glu Glu Gly Met Pro Ala Val Pro His Lys Thr Asn
385                 390                 395                 400

ctc acc tcc ctt aac gtt cca cca atg ggt ctg ttt tat cct aac ctt     1248
Leu Thr Ser Leu Asn Val Pro Pro Met Gly Leu Phe Tyr Pro Asn Leu
                405                 410                 415

ttg gtc cct gaa ata ttt ccc cag cca cca cgt caa tgg ttt caa gac     1296
Leu Val Pro Glu Ile Phe Pro Gln Pro Pro Arg Gln Trp Phe Gln Asp
            420                 425                 430

tac gag aat atg ttg gaa gac act tgg aac atg gac ttt gga ggt ggt     1344
Tyr Glu Asn Met Leu Glu Asp Thr Trp Asn Met Asp Phe Gly Gly Gly
        435                 440                 445

ggt aac atg gga tta cca atg tgg gat agt ttt gca ttt tcg cct tca     1392
Gly Asn Met Gly Leu Pro Met Trp Asp Ser Phe Ala Phe Ser Pro Ser
    450                 455                 460

aaa cca aag aaa gaa gag aag att ggt ctc gct gaa gcc att aca agc     1440
Lys Pro Lys Lys Glu Glu Lys Ile Gly Leu Ala Glu Ala Ile Thr Ser
465                 470                 475                 480

agc att ctc tct gct gga cgc ata gac ctt aga cgg aag ctt ttc tcc     1488
Ser Ile Leu Ser Ala Gly Arg Ile Asp Leu Arg Arg Lys Leu Phe Ser
                485                 490                 495
```

```
agc att caa ttg att ggt ggt gct ggt ttg acg aaa ggt ctt gta gcg      1536
Ser Ile Gln Leu Ile Gly Gly Ala Gly Leu Thr Lys Gly Leu Val Ala
            500                 505                 510

gca gtg gaa gaa aga gtt ctt cac gcg ata cct cca act gaa gcc att      1584
Ala Val Glu Glu Arg Val Leu His Ala Ile Pro Pro Thr Glu Ala Ile
        515                 520                 525

gat aca gtg cag gtt ctg cca tca aga acg gag cca caa ttc gta act      1632
Asp Thr Val Gln Val Leu Pro Ser Arg Thr Glu Pro Gln Phe Val Thr
    530                 535                 540

tgg aaa gga gga gcg ata ttg gga att ctg gac ttt ggg agg gag gct      1680
Trp Lys Gly Gly Ala Ile Leu Gly Ile Leu Asp Phe Gly Arg Glu Ala
545                 550                 555                 560

tgg att gag aga cat caa tgg atg gta aat ggg gtt aat aaa ggt ggt      1728
Trp Ile Glu Arg His Gln Trp Met Val Asn Gly Val Asn Lys Gly Gly
                565                 570                 575

cta aag aag tac aaa gac tct tat cac ctt caa ggt caa gca atg tac      1776
Leu Lys Lys Tyr Lys Asp Ser Tyr His Leu Gln Gly Gln Ala Met Tyr
            580                 585                 590

ttc atc aac ccc tag                                                 1791
Phe Ile Asn Pro
        595
```

<210> 18

<211> 596

<212> PRT

<213> Arabidopsis thaliana

<400> 18

```
Met Ala Lys Pro Lys Ser Asn Ser His Leu Ser Trp Gln Asp Tyr Leu
1               5                   10                  15

Lys Thr Val Ala Pro Thr Gln Ile Leu Ser Glu Arg Gly Ala Asn Leu
            20                  25                  30

Val Val Ile Asn Leu Gly Ser Ala Asn Val Arg Val Gly Leu Ala Met
        35                  40                  45

Asp Glu Lys Pro Phe Asn Val Pro Asn Cys Ile Ala Arg Tyr Ile Thr
    50                  55                  60

Gln Ser Gly Lys Pro Thr Val Val Asp Gln Met Leu Asn Thr Glu Val
65                  70                  75                  80

Thr Thr Asn Gln His Val Asp Arg Glu Arg Ala Tyr Asn Ser Ala Ala
                85                  90                  95
```

```
Ser Leu Leu Lys Ile Leu Phe Leu Asp Glu Ser Ser Ser Ser Gly Ser
            100                 105                 110

Ala Ser Arg Lys Met Gly Arg Ile Asp Gly Tyr Asn Gln Ala Ser Thr
        115                 120                 125

Ile Lys Lys Asp Ser Val Phe Thr Trp Thr Asp Val Tyr Glu Asp Glu
    130                 135                 140

Lys Ile Ser Leu Ala Ser Pro Ala Glu Thr Ser Pro Asp Lys Gly Asp
145                 150                 155                 160

Ala Ser Ala Ser Glu Ala Val Pro Asp Val Thr Asp Ser Lys Asp Thr
                165                 170                 175

Ser Glu Ser Lys Arg Lys Tyr Arg Lys Met Ile Phe Gly Glu Glu Ala
            180                 185                 190

Leu Lys Ile Ser Pro Lys Glu Pro Tyr Cys Leu Tyr His Pro Ile Arg
        195                 200                 205

Arg Gly His Phe Asn Val Ser Pro His Tyr Ser Ala Gln Arg Val Cys
    210                 215                 220

Glu Asp Leu Thr Ala Ile Leu Asp Trp Ile Leu Leu Glu Lys Leu His
225                 230                 235                 240

Ile Thr His Lys Glu Arg Phe Ser Phe His Ala Val Ile Val Val Pro
                245                 250                 255

Glu Thr Phe Asp Thr Arg Glu Ile Lys Glu Met Leu Thr Ile Val Leu
            260                 265                 270

Gly Glu Leu Tyr Phe Asn Ser Ala Val Val His Gln Glu Gly Leu Ser
        275                 280                 285

Ala Val Phe Gly Asn Gly Leu Thr Thr Ala Cys Ile Val Asn Ile Gly
    290                 295                 300

Ala Gln Thr Ser Thr Val Val Cys Val Glu Asp Gly Val Ser Leu Pro
305                 310                 315                 320

Asn Thr Glu Lys Ile Leu Pro Phe Gly Gly Glu Asp Ile Cys Arg Cys
                325                 330                 335

Leu Leu Trp Ile Gln Arg His Tyr Gln Lys Trp Pro Gln Ile Asn Thr
            340                 345                 350

Asp Val Leu Ala Lys Pro Ile Asp Met Leu Met Leu Asn Gln Leu Lys
        355                 360                 365
```

Glu Ser Phe Cys Glu Ile Arg Ala Gly Glu Leu Glu Thr Val Ala Thr
370 375 380

Val His Ser Tyr Glu Glu Gly Met Pro Ala Val Pro His Lys Thr Asn
385 390 395 400

Leu Thr Ser Leu Asn Val Pro Pro Met Gly Leu Phe Tyr Pro Asn Leu
405 410 415

Leu Val Pro Glu Ile Phe Pro Gln Pro Pro Arg Gln Trp Phe Gln Asp
420 425 430

Tyr Glu Asn Met Leu Glu Asp Thr Trp Asn Met Asp Phe Gly Gly Gly
435 440 445

Gly Asn Met Gly Leu Pro Met Trp Asp Ser Phe Ala Phe Ser Pro Ser
450 455 460

Lys Pro Lys Lys Glu Glu Lys Ile Gly Leu Ala Glu Ala Ile Thr Ser
465 470 475 480

Ser Ile Leu Ser Ala Gly Arg Ile Asp Leu Arg Arg Lys Leu Phe Ser
485 490 495

Ser Ile Gln Leu Ile Gly Gly Ala Gly Leu Thr Lys Gly Leu Val Ala
500 505 510

Ala Val Glu Glu Arg Val Leu His Ala Ile Pro Pro Thr Glu Ala Ile
515 520 525

Asp Thr Val Gln Val Leu Pro Ser Arg Thr Glu Pro Gln Phe Val Thr
530 535 540

Trp Lys Gly Gly Ala Ile Leu Gly Ile Leu Asp Phe Gly Arg Glu Ala
545 550 555 560

Trp Ile Glu Arg His Gln Trp Met Val Asn Gly Val Asn Lys Gly Gly
565 570 575

Leu Lys Lys Tyr Lys Asp Ser Tyr His Leu Gln Gly Gln Ala Met Tyr
580 585 590

Phe Ile Asn Pro
595

<210> 19

<211> 29

<212> DNA

<213> Arabidopsis sp.

<400> 19
ctagaagctt ttaaggatta agactctcc 29

<210> 20

<211> 22

<212> DNA

<213> Arabidopsis sp.

<400> 20
ctcgtatgta tccccttct cc 22

<210> 21

<211> 21

<212> DNA

<213> Arabidopsis thaliana

<400> 21
tgatggatct atcaccatca g 21

<210> 22

<211> 21

<212> DNA

<213> Arabidopsis thaliana

<400> 22
ggtgggattc caatcacttt c 21

<210> 23

<211> 21

<212> DNA

<213> Arabidopsis thaliana

<400> 23
tgatggatct atcaccatca g 21

<210> 24

<211> 21

<212> DNA

<213> Arabidopsis thaliana

<400> 24
ggtgggattc caatcacttt c 21

<210> 25

<211> 19

<212> DNA

<213> Arabidopsis thaliana

<400> 25
tttgatgggc caaatgatg 19

<210> 26

<211> 36

<212> DNA

<213> Arabidopsis thaliana

<400> 26
cttccaacct aggtgagatg tttcaacaaa atgtgc 36

<210> 27

<211> 19

<212> DNA

<213> Arabidopsis thaliana

<400> 27
tcaacagcag gacacaagg 19

<210> 28

<211> 37

<212> DNA

<213> Arabidopsis thaliana

<400> 28
cccaatgcct aggaaatccg agttcaacat cctaatc 37

<210> 29

<211> 23

<212> DNA

<213> Arabidopsis thaliana

<400> 29
acaaaccaga tatcagcaca tgg 23

<210> 30

<211> 23

<212> DNA

<213> Arabidopsis thaliana

<400> 30
aacaagtact cgctctcatg ctc 23

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 03/06757

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C07K14/415

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EMBL, BIOSIS, EPO-Internal, SEQUENCE SEARCH, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL/GENBANK/DDBJ 'Online! "Arabidopsis thaliana putative helicase (At3g57300) mRNA, partial cds." retrieved from EBI Database accession no. AY080695 XP002254488 abstract ---	1,2
X	DATABASE EMBL/GENBANK/DDBJ 'Online! 21 March 2002 (2002-03-21) "Arabidopsis thaliana cDNA clone:RAFL09-38-B21, 5'-end." retrieved from EBI Database accession no. AV829055 XP002254489 abstract --- -/--	1-3

[X] Further documents are listed in the continuation of box C. [ ] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
15 September 2003	02/10/2003

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Paresce, D

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 03/06757

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL/GENBANK/DDBJ 'Online! 21 January 2000 (2000-01-21) "Arabidopsis thaliana DNA chromosome 3, BAC clone F2809" retrieved from EBI Database accession no. AL137080 XP002254490 abstract	1,2
X	DATABASE EMBLGENBANK DDBJ 'Online! 1 October 2000 (2000-10-01) "Helicase-like protein F2809.150." retrieved from EBI Database accession no. Q9M2L7 XP002254491 abstract	5
X	SHEN XUETONG ET AL: "A chromatin remodelling complex involved in transcription and DNA processing." NATURE (LONDON), vol. 406, no. 6795, 2000, pages 541-544, XP002254485 ISSN: 0028-0836	6,15
Y	see p. 541, 543-4	1-15
X	EBBERT RONALD ET AL: "The product of the SNF2/SWI2 paralogue INO80 of Saccharomyces cerevisiae required for efficient expression of various yeast structural genes is part of a high-molecular-weight protein complex." MOLECULAR MICROBIOLOGY, vol. 32, no. 4, May 1999 (1999-05), pages 741-751, XP002254486 ISSN: 0950-382X	6,15
Y	see abstract, p. 741-2	1-15
Y	GHERBI HASSEN ET AL: "Homologous recombination in planta is stimulated in the absence of Rad50." EMBO REPORTS, vol. 2, no. 4, April 2001 (2001-04), pages 287-291, XP002254487 April, 2001 ISSN: 1469-221X cited in the application see abstract, p. 287-8	1-15